VOL. XXII FEBRUARY, 1938 No. 1

THE JOURNAL OF APPLIED PSYCHOLOGY

JAMES P. PORTER, *Editor*
OHIO UNIVERSITY ATHENS, OHIO

With a Board of Cooperating Editors
and

PAUL S. ACHILLES HENRY C. LINK
Associate Editors Representing
THE PSYCHOLOGICAL CORPORATION

SOME INTERESTING CONTRIBUTIONS IN THIS ISSUE:

CHRISTIAN A. RUCKMICK. The Truth About the Lie Detector

GLENN E. MITCHELL AND HAROLD E. BURTT. Psychological Factors in the Sales Interview

HARRY R. DESILVA AND RALPH CHANNELL. Driver Clinics in the Field

CARRIE MAE SCOTT. Background and Personal Data as Factors in the Prediction of Scholastic Success in College

PUBLISHED BI-MONTHLY BY THE EDITOR
OHIO UNIVERSITY, ATHENS, OHIO

PRINTED BY THE SCIENCE PRESS PRINTING COMPANY, LANCASTER, PA.

Entered as second-class matter at the post office at Athens, Ohio, under the act of March 3, 1879. Additional entry May 7, 1935, at the post office at Lancaster, Pa. Copyright 1937 by James P. Porter.

The Journal of Applied Psychology

CO-OPERATING EDITORS

W. F. Book, Indiana University, W. V. Bingham, Personnel Research Federation, New York; M. E. Broom, State Teachers College, San Diego, Calif.; H. E. Burtt, Ohio State University; Robert H. Gault, Northwestern University; William Healy, Baker Foundation, Juvenile Court, Boston; H. L. Hollingworth, Columbia University; J. E. W. Wallin, Atlantic University; Herbert S. Langfeld, Princeton University; James H. Leuba, Bryn Mawr College; Henry M. McComas, Johns Hopkins University; Walter D. Scott, Northwestern University; C. E. Seashore, University of Iowa; Daniel Starch, New York; Edward K. Strong, Jr., Leland Stanford Jr. University; L. M. Terman, Leland Stanford Jr. University; E. L. Thorndike, Columbia University; Joseph Jastrow, New School of Social Research, New York; G. M. Whipple, Danvers, Mass.; Helen Thompson Woolley, Teachers College, Columbia University.

The scope of The Journal of Applied Psychology will include the following: (a) The application of psychology to vocational activities, such as law, art, public speaking, industrial and commercial work, and problems of business appeals. (b) Studies of individual mentalities, such as types of character, special talents, genius, and individual differences, including the problems of mental diagnosis and vocational prognosis. (c) The influence of general environment conditions, such as climate, weather, humidity, temperature; also such conditions as nutrition, fatigue, etc. (d) The psychology of everyday activities, such as reading, writing, speaking, singing, playing games or playing musical instruments, sports, etc.

Communications, manuscripts, reviews, summaries, notices, books and offprints for review should be addressed to Dr. James P. Porter, Ohio University, Athens, Ohio. Contributors are requested to follow the conventional method of citation employed by the Journal; in every case the year, volume and pages should be specified.

The Journal of Applied Psychology is published bi-monthly. The subscription price is $6.00 a year. Single numbers $1.25 to $2.00, depending upon the number of pages. Remittances and business communications should be addressed to the Editor, James P. Porter, Ohio University, Athens, Ohio.

A correction: Under date of January 6, 1938, a letter which reads as follows was received from Dr. Charles R. Atwell, one of the authors of the article entitled "Wide Range Multiple Choice Vocabulary Tests" appearing in the October 1937 issue of this Journal:

> ". . . May we invite attention to an unfortunate misprint occurring on page 551, bottom of the second paragraph as fellows:
>
> ' "Preacher" as a mislead for "minister" (minster) depended more upon a misreading of the stimulus word than upon knowing the meaning of the word.'
>
> "The error in question rather helps to emphasize the idea we wish to convey. In reading the galley proof, I overlooked it until it was called to my attention by Dr. Wells, and apparently the printer was unwilling to accept the delete sign we attached."

Vol. XXII FEBRUARY, 1938 No. 1

THE JOURNAL OF APPLIED PSYCHOLOGY

Edited by
JAMES P. PORTER, Ohio University
Athens, Ohio

TABLE OF CONTENTS

JOHN G. JENKINS. Dependability of Psychological Brand Barometers, I. The Problem of Reliability 1

PAUL S. DWYER. An Analysis of Nineteen Occupational Scores on the Strong Vocational Interest Blank 8

GLENN E. MITCHELL AND HAROLD E. BURTT. Psychological Factors in the Sales Interview 17

H. E. GEIGER, H. H. REMMERS AND R. J. GREENLY. Apprentice Attitudes Toward Their Training and the Construction of a Diagnostic Scale 32

CARRIE MAE SCOTT. Background and Personal Data as Factors in the Prediction of Scholastic Success in College 42

CHRISTIAN A. RUCKMICK. The Truth About the Lie Detector 50

HARRY R. DESILVA AND RALPH CHANNELL. Driver Clinics in the Field 59

ELEROY L. STROMBERG. The Relationship of Measures of Visual Acuity and Ametropia to Reading Speed 70

GEORGE J. DUDYCHA. The Estimation of Performance-time in Simple Tasks 79

S. JEAN WOLF. Comparison of the Standard and a "Self-Scoring" Form of the Vineland Adjustment Score Card 87

MARION A. BILLS. Relation of Scores in Strong's Interest Analysis Blanks to Success in Selling Casualty Insurance 97

E. G. WILLIAMSON. A Further Analysis of the Young-Estabrooks Studiousness Scale 105

News and Notes 106

Book Reviews 110

New Books and Pamphlets 116

DEPENDABILITY OF PSYCHOLOGICAL BRAND BAROMETERS

I. THE PROBLEM OF RELIABILITY

JOHN G. JENKINS
University of Maryland

NOTE: In 1936, the Psychological Corporation made five research grants for special research projects in connection with the Psychological Brand Barometers. These grants were made on the basis of outlines for research submitted by a number of different psychologists. The projects selected and carried out were as follows:

Dependability, John G. Jenkins, then at Cornell University
Validity, " " " " " " "
Validity in Respect to Tires, K. W. Oberlin, University of Delaware
Validity Based on Pantry Inventory Method, A. C. Welch, then at University of Minnesota
Comparison with Multiple Choice Method, Herbert Moore, Mount Holyoke College
Validity Based on Pantry Inventory Method, D. B. Lucas, New York University

As a result of Oberlin's study of the automobile tire questionnaire, and the confirmation of these results in Jenkins' study, the question on tires in the Barometers was given up. The present paper is a report on one of these projects and it is hoped that some of the other projects may be reported also.

FOR the past five years psychologists at more than fifty American universities have been contributing their services periodically to the maintenance of a national program of psychological market research. This program, known as the Psychological Brand Barometer, had its inception in a suggestion by Link[1] in 1932 that periodic surveys of buying habits made by psychologically trained investigators at various universities would provide the best possible index to future buying behavior. With the cooperation of psychologists at strategically located universities, the program expanded so

[1] H. C. Link, *The New Psychology of Selling and Advertising*, 1932, 28.

rapidly that in 1935 Link and Likert[2] could speak of a series of reports representing the cooperative labors of sixty psychologists in collecting data on the purchasing habits of more than 42,000 housewives. From the start it was apparent that these data could be utilized to predict trends in buying behavior; and in 1936 the Psychological Corporation issued charts[3] which showed their use in detecting distinct buying-trends.

The Psychological Brand Barometer consists in essence of a series of questions regarding the last purchase of a score of commodities made by the respondent. "These barometers utilize, as measures of distribution, a psychological yardstick, namely the buying habits of comparable groups of customers. Periodically, in nation-wide interviews, people are asked: 'What brand of coffee did you buy last?' 'What soap did you use last for washing dishes?' 'What cigarette are you smoking now?' etc. . . . The Psychological Brand Barometer . . . is not a measure of sales in the ordinary sense. It is a measure of customers. Its unit of measurement is not dollars or pounds but people. Its purpose is to reveal whether *more or fewer* people are buying a given article, and to show such trends in relation to competing articles in the same group."[4]

The practical utility of the method may be regarded as definitely established. For certain commodities, and with proper interpretation of statistical indices, the results already obtained have demonstrated that last-purchase questions can supply the most sensitive indicator of buying-trends thus far developed. No small number of commercial organizations have been sufficiently impressed with the value of the Brand Barometers to pay well for their use. Certain other organizations, it may be remarked, have demonstrated their high appreciation of the method by appropriating it for their own

[2] H. C. Link and R. Likert, Recent scientific techniques in measuring distribution and factors affecting distribution, *International Congress for Scientific Management*, London, 1935, 3.

[3] These charts, privately published in the form of lantern slides, are available to psychologists on application to the Psychological Corporation.

[4] H. C. Link and R. Likert, *op. cit.*, 3.

use, without the benefit of psychologists to administer the questions or to interpret the results.

Granting the practical utility of the method, we are still confronted with the necessity of subjecting it to precisely the same tests of dependability which we should employ in testing any other psychological instrument. We must attempt to discover whether the last-purchase questionnaire actually yields a reliable and valid depiction of our respondents' buying habits or whether the good practical results are due to some less obvious set of conditions. We must find out, in other words, whether people are consistent in naming the last brand of a commodity purchased and whether their statements actually represent their most recent purchases.

MEASUREMENT OF RELIABILITY

The familiar technique of repeating interviews with an identical sampling of respondents was employed in measuring reliability. A group of trained undergraduate interviewers went to 150 assigned B-group homes and filled out there the regular Barometer interview blanks. There were twenty-six questions on this particular blank; nineteen of them were typical last-purchase questions, the remainder being scattered among other types. Our investigation was concerned only with the last-purchase questions. As soon as the interviewers turned in the completed interviews, each sheet was assigned an identifying code-number.

Forty-eight hours after the original interviews, the same interviewers went again to the same homes, asked to interview the original respondent, and, if successful in obtaining the repeat-interview, explained that the original sheet had been lost, and asked the respondent to submit to a second interview. The affability of the respondents was surprising; where we had expected to obtain less than 75 per cent of the repeat interviews, we actually obtained better than 90 per cent. After discarding all sheets which involved a change in respondents between the two interviews, we had, as a result, 100 pairs of

blanks[6] representing the answers of the same interviewees given by the same interviewers under the same general conditions. These pairs were now arranged by code-number and were submitted to an item-by-item analysis in order to determine the nature and the extent of changes made in answers on the repeat-interview.

The results, subdivided by commodities, are given in Table I.

TABLE I
Showing the Percentage of Respondents Naming the Same Brand on Repeat Interview

COMMODITY	PER CENT OF RESPONDENTS NAMING SAME BRAND
Beer (home)	97
Tomato juice	93
Cigarette (women)	93
Toilet soap (men)	91
Face powder	91
Scouring powder	91
Gasoline	91
Coffee	90
Mouthwash	90
Motor oil	90
Beer (away)	90
Toilet soap (women)	89
Cigarette (men)	89
Toothpaste	88
Face cream	88
Flour	88
Shaving soap	88
Hand lotion	87
Tires	85
Average	90.0
A. D.	1.8

Perhaps the most striking thing about these results is the high agreement on reliability for the wide variety of commodities indicated. Even with all of the items included, the A. D. of 1.8 would seem to indicate that the high reliability of the method was relatively independent of the class of

[6] Of the 100 complete interviews finally paired and employed in obtaining our data, 76 were with housewife only, 3 with husband only, and 21 with husband and wife.

product investigated. And when we eliminate the figures for two products for which special determinants are known to be operating, the total range of indices of reliability runs only from 87 per cent to 93 per cent. (The index for beer consumed is out of line because almost all of the respondents denied consumption at home. The index for tires is brought down to a lower figure by the common confusion between "Goodrich" and "Goodyear"; three respondents changed answers from Goodrich to Goodyear while two changed in the other direction.) The conclusion seems inescapable that our sampling in the "B" group was consistent in supplying the name of the brand last purchased, that is, that the last-purchase question ranks high in reliability, at least within the limits of this investigation.

The same data may be used to supply answers to two other questions. The first concerns the variability shown by individual respondents, while the second has to do with the practical outcome of the shifts with regard to the trade-names listed. The first question arises from the fact that it would be theoretically possible to have obtained a gross reliability index of 90 per cent in quite different ways. We might have had 90 interviewees who made no changes in the repeat interview, in combination with 10 who changed every answer; we might have had 100 interviewees who changed 1.9 answers apiece; or we might have had any other combination between these limits. The figures in Table II supply the answer to this query:

TABLE II

Showing the Number of Changes per Respondent on the Total Questionnaire

18	sheets	showed	0	changes
29	"	"	1	"
24	"	"	2	"
15	"	"	3	"
8	"	"	4	"
2	"	"	5	"
2	"	"	6	"
2	"	"	7	"

(there were 19 questions on each sheet)

The results show that only 14 respondents changed more than three answers on their repeat interviews, and this figure must be balanced against the 18 respondents who made no changes. In general the results point to a fairly uniform level of few changes.

Our second question is one of considerable practical significance. When people change their answers, is there any specific drift detectable in these changes? The writer had long secretly espoused a theory that when people first responded to a questionnaire, they would tend to supply well-known trade-names in cases of doubt as to their actual purchases and that, if given a second chance after opportunity to refresh their memories, they would change away from the widely advertised brands. If this theory were true, the group should show a drift away from the better known brands on the repeat interviews. To check the actual nature of the shifts, an analysis was made of the direction of change for each commodity. Obviously it is impossible to reproduce any considerable part of this analysis within the limits of this report. There was, however, such a high degree of consistency in the results that a single illustration may serve to answer the question. This is provided in Table III.

TABLE III

Showing the Changes for Brands of Toothpaste Named on Repeat Interviews

TRADE-NAMES	GAINS	LOSSES	NET CHANGE
Pepsodent	4	3	+1
Colgate	1	2	−1
Squibbs	1	1	0
Listerine	2	1	+1
Ipana	2		+2
Dr. West	1	1	0
Dr. Lyons		1	−1
Prep	1		+1
None and Don't Use		3	−3

The data in this table show two things. First, it will be noted that the net effect of the shifts is small. This is true in the case of all the other questions, many of which show a

considerably larger number of zeros in the 'Net Change' column. When the complete results were shown to the statistical staff of the Psychological Corporation, they were unanimous in affirming that none of the changes was large enough seriously to affect the predictive value of the total. Second, it is apparent that there is no consistent shift toward or away from the more widely advertised brands—a state of affairs which obtains throughout the interview-blank.

In stating that the data point to a reliability of about 90 per cent, we have, of course, failed to give the last-purchase questions the benefit of any allowance for actual changes due to new purchases made in the forty-eight hours intervening between the two interviews. It is quite certain that some percentage of the changes should be credited to new purchases, but there is no trustworthy way of estimating the amount for our sampling. We may assume that the correction for new purchases would be relatively small in the case of tires and possibly quite large in the case of such commodities as cigarettes, coffee, and the like. There is no reason to assume that even the most precise correction for this factor would change the index of reliability sufficiently to have any practical significance. In any event, it would tend to raise, and not to lower, the indicated reliability.

CONCLUSIONS

Using the technique of repeated interview with the identical sampling of respondents, it has been shown that the last-purchase questions employed yielded an index of reliability of 90 per cent. The small range of values for the wide variety of products studied is taken to mean that the obtained index of reliability is relatively independent of the nature of the product. It appears that the essential technique of the Psychological Brand Barometers is not only demonstrated to be of high practical utility, but that it is also shown to be high in measured reliability. The problem of validity will be discussed in a later paper.

AN ANALYSIS OF 19 OCCUPATIONAL SCORES OF THE STRONG VOCATIONAL INTEREST TEST GIVEN TO 418 STUDENTS ENTERING THE UNIVERSITY OF MICHIGAN MEDICAL SCHOOL DURING THE YEARS 1928, 1929, 1930

PAUL S. DWYER
University of Michigan

DURING the years 1928, 1929, and 1930 the Strong Vocational Interest Test was given to the applicants for admission to the Medical School of the University of Michigan. Of these applicants, 418 matriculated. The occupational scores have been computed for these 418 cases and a table of correlations prepared.

In computing correlations, the score on each occupation was coded to a single digit so that one column of a Hollerith card would be sufficient for each occupation. The resulting correlations, given to two decimal places, were not corrected for grouping. The names of the 19 occupations scored, and the correlations between the scores, are presented in Table I.

Two previous tables of correlations, though for slightly different groups of occupations, have been published: (a) the Strong results for college seniors (1:200), hereafter referred to as the college results, (b) the Carter, Pyles, Bretnall (2:85) results for high-school students, hereafter referred to as the high-school results. One naturally wonders how these results agree with the results of Table I (which are to be called the pre-medical results). Carter (2) has compared the correlations between the college results and the high-school results and has found a correlation of $r = .77$. The equivalent correlation

TABLE I

The Intercorrelation of the 19 Occupations

NO.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Occupation	Dr.	L.	En.	T.	Phy.	C.	M.	J.	A.	Psy.	Min.	Y.M.	F.	Ad.	C.P.A.	L.I.	P.	P.A.	R.E.
1 Dr.		.41	.61	−.11	.77	.54	.69	.45	.69	.51	−.10	−.47	.30	.18	.17	−.33	−.46	−.23	−.10
2 Law.	.41		−.04	−.35	.14	−.21	.27	.81	.18	.09	−.03	−.44	−.25	.70	.44	.28	−.38	−.22	.50
3 Eng.	.61	−.04		−.01	.78	.73	.61	−.11	.51	.38	−.20	−.37	.54	−.23	.12	−.49	−.16	.21	−.27
4 Teacher	−.11	−.35	−.01		.09	.27	.16	−.43	.01	.43	.83	.68	−.04	−.56	.01	−.26	.65	−.22	−.69
5 Physicist	.77	.14	.78	.09		.78	.89	.17	.67	.71	−.06	−.39	.37	−.06	.19	−.65	−.33	−.22	−.40
6 Chemist	.54	−.21	.73	.27	.78		.63	−.18	.54	.70	.10	−.17	.47	−.35	−.04	−.76	−.02	−.18	−.61
7 Math.	.69	.27	.61	.16	.89	.63		.29	.63	.72	−.01	−.34	.28	.02	.30	−.56	−.30	−.28	−.36
8 Jour.	.45	.81	−.11	−.43	.17	−.18	.29		.32	.17	−.27	−.46	−.27	.81	.30	.15	−.49	−.37	.43
9 Arch.	.69	.18	.51	.01	.67	.54	.63	.32		.61	.01	−.31	.22	.14	.05	−.48	−.27	−.35	−.27
10 Psychol.	.51	.09	.38	.43	.71	.70	.72	.17	.61		.32	−.04	.13	−.07	.09	−.73	.02	−.58	−.60
11 Minister	−.10	−.03	−.20	.83	−.06	.10	−.01	−.27	.01	.32		.74	−.25	−.40	−.02	−.11	.59	−.36	−.61
12 Y.M.C.A.	−.47	−.44	−.37	.68	−.39	−.17	−.34	−.46	−.31	−.04	.74		−.18	−.47	−.15	.10	.68	−.05	−.42
13 Farmer	.30	−.25	.54	−.04	.37	.47	.28	−.27	.22	.13	−.25	−.18		−.41	−.35	−.40	−.15	.22	−.21
14 Advert.	.18	.70	−.23	−.56	−.06	−.35	.02	.81	.14	−.07	−.40	−.47	−.41		.24	.31	−.42	−.25	.63
15 C.P.A.	.17	.44	.12	.01	.19	−.04	.30	.30	.05	.09	−.02	−.15	−.35	.24		.10	−.03	.11	.17
16 Life Ins.	−.33	.28	−.49	−.26	−.65	−.76	−.56	.15	−.48	−.73	−.11	.10	−.40	.31	.10		.03	.31	.70
17 Personnel	−.46	−.38	−.16	.65	−.33	−.02	−.30	−.49	−.27	.02	.59	.68	−.15	−.42	−.03	.03		.07	−.36
18 Pur. Agent	−.23	−.22	.21	−.22	−.22	−.18	−.28	−.37	−.35	−.58	−.36	−.05	.22	−.25	.11	.31	.07		.30
19 Real Estate	−.10	.50	−.27	−.69	−.40	−.61	−.36	.43	−.27	−.60	−.61	−.42	−.21	.63	.17	.70	−.36	.30	

for the occupations common to the three studies is r = .79. In the same manner, correlation between the high-school results and the pre-medical results for occupations common to the three studies is r = .94, while the correlation between college results and pre-medical results is r = .89. The size of the coefficients in the case of the pre-medical group makes it appear that the selection of a group of medical students has not violently disturbed the general correlations and that the pre-medical group is the one which may best serve as a norm for the three groups. This argument is further strengthened by the fact that the pre-medical group is based on 418 cases, while the high-school group is based on 133 cases and the college group on 287 cases.

Table I was then subjected to the first steps of a multiple factor analysis. The centroid method revealed that four factors produced the correlations to within small errors. The centroid solution is given in Table II. The columns on the right give the communality, $h_j^2 = a_{j1}^2 + a_{j2}^2 + a_{j3}^2 + a_{j4}^2$ and h_j.

TABLE II
Centroid Results

TEST	OCCUPATION	I	II	III	IV	h^2_j	h_j
1	Doctor	.830	.079	.082	−.205	.74	.86
2	Lawyer	.420	−.565	.490	−.286	.82	.91
3	Engineer	.628	.345	−.411	−.402	.84	.92
4	Teacher	−.247	.774	.426	−.202	.88	.94
5	Physicist	.869	.384	−.057	−.183	.94	.98
6	Chemist	.606	.652	−.205	−.054	.84	.92
7	Mathematician	.833	.306	.157	−.216	.86	.93
8	Journalist	.527	−.602	.496	.116	.90	.95
9	Architect	.727	.246	.130	.086	.61	.78
10	Psychologist	.626	.565	.371	.146	.87	.93
11	Minister	−.297	.610	.626	−.119	.87	.93
12	Y.M.C.A.	−.659	.517	.341	.011	.82	.90
13	Farmer	.272	.336	−.565	−.043	.51	.71
14	Advertiser	.311	−.744	.353	.186	.81	.90
15	C.P.A.	.185	−.221	.290	−.421	.35	.58
16	Life Insurance	−.536	−.632	.140	−.288	.79	.89
17	Personnel	−.561	.481	.217	−.132	.61	.78
18	Pur. Agent	−.292	−.161	−.543	−.474	.63	.79
19	Real Estate	−.075	−.889	−.103	−.140	.83	.91

The centroid axes must be rotated to demonstrate simple structure before psychological meaning can be given to the

factors. There is some information, however, which is definitely determined by the centroid solution; first, the table of correlations may be obtained, aside from small residuals, from four common factors and, second, the communality is definitely determined since it is invariant under rotation of axes.

Now the multiple correlation of any occupation j with the n – 1 remaining occupations is equal to, or less than, the square root of the communality of occupation j as Roff (3:4) has shown. Hence the last column of Table II shows for each occupation a value which the multiple correlation coefficient may not exceed even though the occupation score is predicted from all the n – 1 occupations. It would be very useful if we could predict each occupation from four "fundamental" occupations rather than from the n – 1 remaining ones. This is possible, from a practical standpoint, if the multiple correlation of a given occupation with the four "fundamental" occupations is not much less than the multiple correlation of that occupation with the n – 1 remaining occupations. Any multiple correlation cannot be greater than the square root of the communality, given in the last column of Table II.

The next step is the selection of the "fundamental" occupations. If simple structure is present and if each primary trait has one of the occupations very close to it, it is possible to substitute the actual occupation for the primary trait with little loss of predictive power.

Formally this process calls for the isolation of the primary traits, the demonstration of simple structure, and then the selection of appropriate occupations. In practice, however, it may be possible to make estimates of the position of primary traits and to select "fundamental" occupations from a graphical representation of the centroid solution. The merit of the selection is determined by the extent to which the multiple correlations, obtained by predicting from the "fundamental" occupations, approach the last column of Table II.

Using reference vectors I, II, III of Table II as axes, the augmented coordinates of the different occupations were

plotted. The resulting configuration, with consideration of reference vector IV, led to the selection of the occupations, physicist (5), journalist (8), minister (11), and life insurance salesman (16) as "fundamental" occupations. There is no proof here that these tests are the best selections. It is probable that there are other combinations which would serve as well. The justification for this selection lies in the fact that, for most of the occupations, the multiple correlation coefficient predicted from these four fundamental occupations, as shown in the next to the last column of Table III, is not much less than the square root of the communality.

The normal equations then become

$$\begin{aligned}
\beta_1 + .17\,\beta_2 - .06\,\beta_3 - .65\,\beta_4 &= r_{k,5} \\
.17\,\beta_1 + \beta_2 - .27\,\beta_3 + .15\,\beta_4 &= r_{k,8} \\
-.06\,\beta_1 - .27\,\beta_2 + \beta_3 - .11\,\beta_4 &= r_{k,11} \\
-.65\,\beta_1 + .15\,\beta_2 - .11\,\beta_3 + \beta_4 &= r_{k,16}
\end{aligned}$$

in which the $r_{k,j}$ may be inserted after manipulative solution is complete. Solution may be made by determinants or by an adaptation of the Doolittle method (4). In any case

$$\begin{aligned}
\beta_1 &= 2.000r_{k,5} - .512r_{k,8} + .135r_{k,11} + 1.391r_{k,16}. \\
\beta_2 &= -.512r_{k,5} + 1.227r_{k,8} + .247r_{k,11} - .490r_{k,16}. \\
\beta_3 &= .135r_{k,5} + .247r_{k,8} + 1.093r_{k,11} - .171r_{k,16}. \\
\beta_4 &= 1.391r_{k,5} - .490r_{k,8} + .171r_{k,11} - 1.997r_{k,16}.
\end{aligned}$$

The values of $r_{k,5}$, $r_{k,8}$, $r_{k,11}$ and $r_{k,16}$ are taken from Table I and the β's computed for each test k. The results are given in Table III.

The value of the multiple correlation coefficient is found by the formula

$$r_{k\,.\,5,8,11,16} = \sqrt{\beta_1 r_{k,5} + \beta_2 r_{k,8} + \beta_3 r_{k,11} + \beta_4 r_{k,16}}.$$

It appears that there has not been much loss of predictive power through the use of the four occupations chosen as "fundamental" although the selection of other occupations might have improved the correlations slightly.

Table III also reveals an approximation to simple structure. It is desired that at least one of the β's be zero in the expression

TABLE III

Values of the Regression Coefficients

TEST	OCCUPATION	β_1	β_2	β_3	β_4	r	h
1	Doctor	.837	.295	*.049*	*.175*	.85	.86
2	Lawyer	.251	.778	.234	.352	.87	.91
3	Engineer	.907	−.344	−.224	*.127*	.85	.92
4	Teacher	*.150*	−.242	.769	−*.041*	.88	.94
5	Physicist	1.000	*.000*	*.000*	*.000*	1.00	.98
6	Chemist	.608	−.223	*.040*	−.327	.86	.92
7	Mathematician	.851	*.172*	*.085*	−*.024*	.91	.93
8	Journalist	*.000*	1.000	*.000*	*.000*	1.00	.95
9	Architect	.509	.287	*.098*	−*.181*	.72	.78
10	Psychologist	.360	.281	.363	−.498	.88	.93
11	Minister	*.000*	*.000*	1.000	.000	1.00	.93
12	Y.M.C.A.	−.305	−.231	.660	*.009*	.85	.90
13	Farmer	.288	−.387	−.358	−*.194*	.62	.71
14	Advertiser	−*.157*	.774	−*.193*	*.070*	.86	.90
15	C.P.A.	.414	*.094*	*.070*	.363	.36	.58
16	Life Insurance	*.000*	*.000*	*.000*	1.000	1.00	.89
17	Personnel	−.279	−.301	.485	−*.058*	.72	.78
18	Purchasing Agt.	*.132*	−.582	−.462	.433	.70	.79
19	Real Estate	−*.128*	.239	−.495	.526	.91	.91

for each occupation. In addition to the usual variations which make these weights small, rather than zero, we have additional variations due to the fact that the occupations chosen are not the precise reference vectors needed to establish simple structure. That simple structure is approximately attained is evident from an examination of those β's whose absolute value is less than .20. All the occupations, with the exception of lawyer and psychologist, have one or more weights less than .20 in absolute value. Most of them have at least one less than .10 in absolute value.

In predicting the scores it is preferable to use actual, rather than standard, units. The regression equations then take the form

$$S_k = b_0 + b_1 S_5 + b_2 S_8 + b_3 S_{11} + b_4 S_{16},$$

where the S's represent the actual scores. The values of the b's for the different occupations may be computed from the betas, means, and sigmas. They may also be computed from the actual ΣS_i, $\Sigma S_i S_j$ which were used in figuring the values

of the correlation coefficients. The value of the b's, computed by the latter method to three decimal places, are given in Table IV.

TABLE IV

The Values of the b Coefficients

TEST	OCCUPATION	b_0	b_1	b_2	b_3	b_4
1	Doctor	227	.474	.197	.029	.161
2	Lawyer	138	.122	.488	−.017	.304
3	Engineer	376	.538	−.249	−.130	.128
4	Teacher	399	.104	−.190	.481	−.037
5	Physicist	000	1.000	.000	.000	.000
6	Chemist	654	.337	−.153	.023	−.323
7	Mathematician	−58	.774	.180	.080	−.036
8	Journalist	000	.000	1.000	.000	.000
9	Architect	357	.335	.223	.066	−.226
10	Psychologist	470	.283	.239	.269	−.623
11	Minister	000	.000	.000	1.000	.000
12	Y.M.C.A.	429	−.296	−.280	.643	.023
13	Farmer	767	.197	−.305	−.225	−.205
14	Advertiser	161	−.137	.797	−.164	.106
15	C.P.A.	265	.140	.094	.037	.211
16	Life Ins. Sales	000	.000	.000	.000	1.000
17	Personnel Manager	587	−.126	−.178	.228	−.030
18	Purchasing Agt.	538	.064	−.323	−.205	.350
19	Real Estate Sales	302	−.082	.173	−.304	.569

The b's in Table IV are applicable to the coded scores. These may be obtained from the Hollerith scores (weights range from 0 to 60) by subtracting 12000. They may be obtained from results using scoring keys of range −30 to 30 by adding 600.

As an illustration we may take the scores on the four "predicting" occupations (for an individual) and compute his grade on vocation 1, doctor. The Hollerith scores are

S_5' (Physicist score) = 12591
S_8' (Journalism score) = 12467
S_{11}' (Ministry score) = 12701
S_{16}' (Life Insurance score) = 12381

so that

$S_5 = 591$, $S_8 = 467$, $S_{11} = 701$, $S_{16} = 381$ and

Predicted $S_1 = 227 + (.474)(591) + (.197)(467) + (.029)(701) + (.161)(381)$

$= 681$

Predicted $S_1^1 = 12681$. The actual S_1^1 in this case was 12683.

If the -30 to 30 range is used we would have

$S_4^{11} = S_4^1 - (420)\ (30)$, $S_4^{11} = -9$,

$S_8^{11} = -133$, $S_{11}^{11} = 101$, $S_{16}^{11} = -219$.

Adding 600 we get the values S_5, S_8, S_{11}, S_{16}, and the resulting $S_1 = 681$. The predicted S_1^{11} is then 81 as compared with the actual 83.

SUMMARY

1. The pre-medical correlations between occupational scores were very similar to the high-school correlations and the college correlations.

2. A centroid solution of the pre-medical correlations revealed four factors and, for most occupations, large communalities.

3. Interpretation of the centroid solution permitted the tentative selection of four of the occupations as "fundamental."

4. Each occupation was predicted from the four "fundamental" ones. The resulting multiple correlation coefficients were not much less than correlation coefficients which would be obtained by predicting each occupation from all the remaining occupations. Exceptions were the occupations of C.P.A. and farmer.

5. It was possible to predict the majority of the occupations with a multiple correlation of .80 or better. The occupations, farmer and C.P.A., alone, resulted in a multiple correlation below .70.

6. Multipliers are provided by which scores on the different occupations may be approximately obtained from the scores on the occupations (a) physicist, (b) journalist, (c) minister, (d) life insurance salesman.

ACKNOWLEDGMENT

This study has been carried on by the office of Educational Investigations of the University of Michigan, under the general

direction of C. S. Yoakum. Special acknowledgment should be made to Mrs. Charlotte Horner who supervised the computation of the table of correlations and to the National Youth Administration which provided assistants.

REFERENCES

1. THURSTONE, L. L. "A Multiple Factor Study of Vocational Interests," *Personnel J.*, 1931–32, 10, 198–205.
2. CARTER, H. D., PYLES, M. K., and BRETNALL, E. P. "A Comparative Study of Factors in Vocational Interest Scores of High School Boys," *J. Educ. Psychol.*, Feb., 1935, 26, 81–98.
3. ROFF, MERRILL. "Some Properties of the Communality in Multiple Factor Theory," *Psychometrika*, June, 1936, 1, 1–6.
4. DWYER, P. S. "The Simultaneous Solution of Groups of Regression Equations and Associated Multiple Correlation Coefficients," *Ann. Math. Statist.*, December, 1937.

PSYCHOLOGICAL FACTORS IN THE SALES INTERVIEW

GLENN E. MITCHELL AND HAROLD E. BURTT
Ohio State University

INTRODUCTION

EXPERIMENTAL work on problems of direct selling has lagged behind that devoted to advertising, doubtless because of the greater difficulty of setting up an experimental situation with objective material in the former case. One effort along this line has recently been reported.[1] The method consisted essentially of enacting a sales interview before a group of persons who evaluated it in various respects. In the experiment to be reported the interview was rated by both the "salesman" and the "prospect" (experimenter and subject).

PROBLEM

The sales interview is characterized by various methods of approach such as breezy, dignified or domineering and by supplementary techniques such as a demonstration. These methods are designed to produce certain changes in the prospect such as arousing interest or confidence or persuading him to purchase. The experiment employed a few such methods and determined the reactions of the subject in response to them. Four pairs of contrasting methods were studied: (1) a demonstration versus an oral elaboration; (2) presentation of facts versus short circuit appeals; (3) a breezy versus a dignified approach; (4) a domineering versus a friendly approach.

The problem was limited to the comparative effect of the two elements of each pair upon the subject. The latter estimated

[1] McKinney, Fred. An Empirical Method of Analyzing a Sales Interview. *Journal of Applied Psychology*, 1937, 21, 280–299.

each sales talk with reference to his interest, feeling, confidence, willingness to accept the proposition, the salesman's attitude and the persuasiveness of the talk. The experimenter likewise made judgments regarding the subject's interest, the extent to which he engaged in conversation, the nature of the questions he asked, his inspection of the product if it was offered and his posture. The above selling procedures and the above items of judgment were not presumed to be exhaustive but were chosen as techniques that are frequently employed in selling and as items which presumably could be rated with some reliability in a single interview.

LIMITATIONS OF THE EXPERIMENT

This research should be considered as a contribution to technique rather than a final determination of the effectiveness of certain sales procedures. The results are subject to the sampling error and the fact that only one experimenter (salesman) was involved. While some significant differences were obtained, they indicate merely that a certain selling procedure is superior to another for *this* salesman. However, such trends are worth noting and suggest the advisability of similar research with other experimenters.

Studies of this sort also have another limitation, *viz.*, the extent to which the laboratory situation is dissimilar to the real life situation with respect to psychological factors. This is the same limitation that characterizes almost any laboratory experiment in social psychology where there is an obvious element of artificiality and the subject is called upon to imagine himself in a real situation. While the same mechanisms such as suggestion or emotion are involved in both cases there is a difference in degree. The only final answer is to check the laboratory results by subsequent field experiments. This may be possible even in problems like the present. Preliminary work, for instance, was carried on by a student in our laboratory who was actually selling bonds.[2] Prior to each interview,

[2] Landis, M. H. Psychological Factors in the Sales Approach. M. A. Thesis. Ohio State University, 1926.

he determined by random procedure which of several approaches would be used in the first few minutes of the interview. He then employed a rating scale which he filled out immediately after the interview evaluating the reactions of the prospects on specific points. A procedure of this sort conceivably could be extended to include a large number of variables.

MATERIALS

The materials consisted of four pairs of sales talks and two rating scales. The two members of each pair of talks were equated as far as possible. They involved the same product but with a different brand name. They were of approximately the same length. The salesman asked the same number of questions at approximately the same points in the talks.

The first pair involved coffee and compared a demonstration with a merely oral elaboration of the product. The demonstration employed a bottle of large coffee beans to show the type of beans sold, a bottle of small beans for an unfavorable comparison, a sales chart indicating the increasing popularity of this brand and a bottle of ground coffee to demonstrate the aroma. The following is an excerpt from this talk:

"Here is a fine coffee that is very economical. . . . The fact is you get eight more cups of coffee per pound when you buy Mikado coffee. That is because the hull and chaff are removed and only the large, select coffee beans are used. Look at those coffee beans. See how large and clean they are—no chaff, no hull. . . . Now look at the kind of coffee beans that are ground into ordinary coffee. See the small straggly looking beans. Here, take one of these in each hand and look at the difference. . . . Mikado is known everywhere for its flavor. Just smell the aroma of this freshly ground Mikado coffee."

The corresponding talk minus the demonstration included the following: "When you try Dow coffee, you will certainly say that it is the finest coffee you have ever tasted. More than that, you will get eight extra cups from every pound of it. You get this extra coffee because no chaff or hull is ground into Dow

coffee. Then too only the large full-grown coffee beans are used. . . . Everywhere Dow coffee is known for its flavor. The aroma of freshly ground Dow is tops."

The second pair of sales talks included a presentation of facts or a long-circuit approach in comparison with a more short-circuit technique—both applied to a radio receiver. The former involved a maximum of information such as compactness, type of circuit and number of tubes, while the latter included a minimum of factual material and stressed such things as being up-to-date or starting the day off with music.

The third comparison dealt with a breezy versus a dignified approach in selling toilet soap. In each instance, an actual bar of soap was presented. The following are excerpts from the breezy informal talk. "Good morning, good morning. How is your supply of toilet soap. . . . Mary Lou is a fine soap for the complexion and bath. You can just put a couple of bars in the bathroom and let the kids wash their necks and ears with it. You will find that they will like it because it is mild and sudsy. You will like it because it is fresh and clean and your old man will like it because it lasts and lasts." The other soap talk was more dignified and mentioned such things as the hard-milled feature, the aroma, and its value to the whole family.

For the comparison between a domineering and a friendly approach, the main difference was in the delivery rather than in the content of the sales talk. In the former, the salesman stood up, talked more loudly and used gestures. At the conclusion of the talk, the subject was offered a blank and a pencil with which to sign an order for the product. The corresponding talk omitted these procedures. Both talks dealt with a universal book of facts under two fictitious names and employed essentially long circuit methods.

The graphic rating scales executed by the subject and the experimenter are shown in Figure 1. These were used in mimeographed form. The names attached to these items for purposes of discussion appear at the left of Table I in the same

Subject's Scale

Interview was absorbing. I did not want to miss a single word.	I was interested in the interview. I wanted to get the ideas presented.	Interview was mildly interesting throughout.	Interview was often uninteresting. Listening required some effort.	I was bored by the interview. I was glad when it was over.
The whole experience was very disagreeable.	Interview was rather unpleasant for the most part.	Interview aroused no particular feeling. It was neither pleasant nor unpleasant.	Interview was moderately pleasant.	Interview was very enjoyable. I would like other similar interviews.
I felt complete confidence in the salesman. I was willing to accept anything he said.	I was inclined to trust the salesman. He was very fair.	I felt neither confidence nor distrust.	I was rather suspicious of the salesman. I felt he might be "putting one over."	I distrusted the salesman. I felt on the defensive throughout.
I accepted the proposition willingly. It seemed reasonable and attractive.	I accepted the proposition without much hesitation.	It was difficult to make a decision. Uncertain as to whether I should accept.	I accepted the proposition with some misgivings.	I felt the proposition was forced upon me against my better judgment.
Salesman was very timid. He seemed to expect a refusal.	Salesman was timid. Did not try to control the situation.	Salesman was hesitant and lacked enthusiasm.	Salesman controlled the situation but not to an objectionable degree.	The salesman was domineering and commanding at all times.
I would not use this brand under any consideration.	I would prefer some other brand to this one.	I would not object to using this brand.	I would consider this brand carefully before buying any other.	I would buy this brand at once.

FIGURE 1

Experimenter's Scale

Listened very intently to every word.	Showed a definite interest.	Was merely respectful. Rather passing interest.	Was apparently more interested in something else.	Inattentive to the extent of discourtesy.
Criticized and objected at every opportunity.	Raised some objections.	Spoke with reluctance.	Primarily "Yes" and "No" replies.	Details asked for. Intelligent comments.
Intelligent questions—details, cost, etc.	General queries only.	Questions with obvious answers.	A few questions. Hesitatingly and irrelevant.	No questions.
Went through a complete operation or inspection of the product.	Examined product and followed explanations.	Held the product and looked at it rather casually.	Hesitated in taking product. Hardly looked at it.	Refused to take the product.
Relaxed, half asleep. Never looked at salesman or product.	Restless, looking elsewhere most of the time.	Looked at salesman or product with any special stimulus to attention.	Looked away and relaxed occasionally.	Watched salesman or product continuously. Upright or leaning forward.

FIG. 1. *(Continued)*

order in which the items occur in the scale. The phrases on the scale were staggered so that a favorable rating would sometimes be at the left and sometimes at the right end, in the conventional effort to minimize the halo effect.

The subjects were employed one at a time and an effort made to approximate the conditions in an actual sales situation. The instructions requested the subject to assume that he was in the market for the product in question although not necessarily for that particular brand and he was encouraged to participate in the interview by interrupting with questions or criticisms. This participation of the subject in the interview was furthered by the use of questions by the experimenter as part of the continuity. These occurred at fixed points and each was followed by a pause. The omission of information regarding prices, terms and guarantees was calculated to give rise to queries. Replies which were made to the spontaneous questions asked by the subjects were standardizd so far as possible. This was feasible in most cases because of the similarity of the questions asked by the different subjects.

After each talk which occupied approximately a minute and a half, the subject and the experimenter filled out their respective rating scales simultaneously. The eight talks were given in a single experimental period.

The order was varied systematically in order to average out errors due to primacy and recency. The talks were numbered from 1 to 8. The members of a given pair were always separated by three intervening talks. Some subjects had the talks in the order 1, 2, 3, 4, 5, 6, 7, 8. Others had them in the order 2, 3, 4, 5, 6, 7, 8, 1 and so on for six other sequences. Of the forty subjects who participated, five had each of these sequences.

The subjects consisted of forty students at the Ohio State University, twenty-six male and fourteen female. One of the writers (Mitchell) served as experimenter throughout. He had considerable previous experience in personal salesmanship both in retail stores and from door-to-door.

The graphic rating scales were scored by means of a ten-division stencil thus converting the actual mark made by the subject or experimenter into a number from 1 to 10, where 10 reflected most favorably upon the selling procedure. The raw data were recorded in terms of these ten possible scores for each variable such as interest, together with other notations as to the position of that particular talk in the series presented to that individual subject.

RESULTS

The principal results dealing with the comparison between the two members of each pair are presented in Table I. Data for the four comparisons are given in separate blocks of the table. The column at the left indicates the aspect which was rated. In each block the first two columns give the mean rating of the 40 subjects for the two types of approach or selling technique indicated at the head of the column. A maximum score of 10 was possible. The following column gives the difference between the two means expressed as a percent of the smaller and the last column gives the critical ratio, that is, the difference between the means divided by its standard deviation. The upper portion of the table gives the ratings made by the subject and the lower portion the ratings made by the experimenter. For example, when a demonstration in selling coffee is compared with a verbal elaboration, the subjects when rating their interest in the sales talk, yielded an average rating of 7.8 for the former and 6.0 for the latter. This represents a difference of 30% which is statistically significant as indicated by the critical ratio of 4.4. Similarly when the experimenter was rating the apparent interest of the subject, the average of his 40 ratings was 9.5 for demonstration and 8.9 for elaboration.

Considering the first comparison, viz., demonstration versus elaboration, it is obvious from the table that the former is consistently superior. In every respect the subject rates the demonstration more highly on the average, the differences ranging from nineteen to thirty-five percent and all being statistically significant. In other words, the subjects found

the demonstration more interesting and pleasant, arousing greater confidence and a more pronounced tendency to accept the proposition, heightening the impression that the salesman maintained control of the situation and involving a greater persuasiveness. Furthermore the demonstration produced more conversation and more questions on the part of the subject, with rather large and significant differences. There was a small but significant difference in the apparent interest of the subject. This significance is somewhat due to the small variability of the experimenter's ratings on this point. The subject's posture is of doubtful significance. The item of inspection is not included in this comparison inasmuch as there was nothing to inspect in the "elaboration" interview.

Turning to the next comparison between the presentation of facts (long circuit) and a mere short-circuit appeal, the results, on the whole, favor the former. The differences are especially marked with reference to the prospect's confidence and his acceptance of the proposition. The long circuit is far more favorable from this standpoint. The ratings as to the salesman's attitude, however, are not significantly different for the two techniques. The experimenter had some suspicion that the large differences may have been partially attributable to rationalization on the part of the subject, and the tendency to rate the factual approach more highly because it is more commendable to govern one's behavior by such considerations rather than by feelings and emotions. The ratings by the experimenter showed no significant differences between the two techniques except a greater amount of conversation on the part of the subjects in the factual approach. A qualitative point not revealed by the rating scale but noted by the experimenter was a greater number of objections raised by the subject in the case of the short circuit approach.

With reference to the breezy versus the dignified approach, the results are less striking. The negative percentages in the table indicate the superiority of the dignified talk. Evidently the dignified approach produces a greater tendency to accept

the proposal. The breezy approach yields a judgment that the salesman was more in control of the situation and the subject participated more in the conversation. Other differences are not significant. The original data show larger standard deviations of the ratings by the subjects than occur in most of the other sales talks. It appears that some people react favorably to the breezy approach whereas others react antagonistically to it. This would seem to be a point on which the salesman would have to adapt himself to the individual.

Finally, with reference to the domineering versus the friendly approach, the results consistently favor the latter by large and significant differences in the subject's ratings. The item regarding the salesman's attitude is not very pertinent and merely indicates that the subjects did detect the domineering aspect, and marked their rating scale accordingly. The experimenter's ratings are not entirely significant but the larger ones also favor the friendly approach.

The data were analyzed from another standpoint by determining for each item on the rating scale the percentage of the subjects expressing greater interest, confidence, etc., for a particular talk of a comparison. This analysis does not add anything to that already presented in Table I and will not be discussed. On the whole, it corroborates the findings in that Table.

Certain minor aspects of the data deserve brief mention. The question arose as to the correlation between the subject's estimate of interest and the experimenter's estimate. The practical question would be to what extent the salesman can judge the interest of the prospect by observing his external behavior. Such correlations were computed in the present case for each kind of sales talk. The correlations range from .30 to $-.14$ and average .14. With forty cases in each example, these correlations are not very significant and with one exception are less than three times their probable error. This result if corroborated in further experiments, is rather disquieting as indicating difficulty in judging a prospect's interest. Final con-

TABLE I

SUBJECT'S RATINGS	DEMONSTRATION	ELABORATION	% DIFFERENCE	CRITICAL RATIO	FACTS	APPEALS	% DIFFERENCE	CRITICAL RATIO	BREEZY	DIGNIFIED	% DIFFERENCE	CRITICAL RATIO	FRIENDLY	DOMINEERING	% DIFFERENCE	CRITICAL RATIO
Interest	7.8	6.0	30	4.4	7.2	5.6	28	4.0	6.8	6.1	11	1.4	6.5	4.2	55	4.5
Feeling	7.9	6.1	29	4.5	7.6	5.8	31	4.1	6.8	6.6	3	.4	6.4	4.1	56	4.8
Confidence	7.4	6.1	21	3.2	7.5	4.7	59	7.0	5.5	6.6	−20	2.3	5.9	2.6	127	9.7
Proposition	7.7	5.7	35	4.0	7.2	4.8	50	4.8	4.8	6.1	−27	3.2	5.3	2.0	166	8.0
Salesman's Attitude	7.4	6.2	19	3.8	7.2	6.7	7	1.3	8.2	6.5	26	4.5	6.7	9.6	−43	10.3
Persuasiveness	7.2	5.7	26	4.0	6.7	4.6	45	5.0	5.0	5.3	−6	.6	5.6	3.3	70	6.0
EXPERIMENTER'S RATINGS																
Interest	9.5	8.9	7	3.5	9.1	9.0	1	.4	9.3	8.8	5	1.2	8.6	9.1	−6	2.5
Conversation	7.6	5.9	29	4.0	7.5	5.7	31	4.0	6.8	5.4	26	3.1	6.7	5.3	26	2.9
Questions	5.9	4.1	44	3.0	4.4	4.0	10	.7	4.6	5.0	−9	.7	5.7	3.5	63	2.8
Inspection									8.1	8.0	1	.5	8.0	7.3	10	2.2
Posture	8.9	8.6	3	2.3	8.8	8.8	0	0	8.9	8.5	5	2.6	8.4	8.8	−5	2.2

clusions should not be drawn until the experiment has been tried with other salesmen.

Further study was made of the correlations between the subject's estimated interest and the other ratings made by the experimenter. Many of these correlations were computed and ranged from .16 to – .19. Evidently the subject's posture, conversation or questions as judged by the experimenter do not give any closer indication of the subject's actual interest than does the experimenter's direct estimate of such interest.

Further study was made of the inter-correlations between the various ratings made by the subject. These were computed separately for each of the eight sales talks. The data are not worth presenting in detail. Suffice it that almost all the correlations were above .50 except when the item "salesman's attitude" was involved. Evidently the effort to avoid the "halo" effect by staggering the items on the graphic scale was not entirely successful.

Some consideration was given to the order of presentation of the various talks with reference to primacy and also the temporal relation of a talk to the other one for the same product. It is apparent that some of the judgments are affected by having had a previous similar talk by way of comparison. For instance, having first had the factual talk about the radio receiver, the subject was then a little less skeptical regarding the short circuit sales talk for a radio receiver. The breezy talk for toilet soap received more favorable ratings when it preceded rather than followed the more dignified talk for the same product. It was apparent that such factors were playing a rôle but the data are not extensive enough to warrant conclusions. They suggest that the salesman's success might depend somewhat on what other salesmen had been working on the prospect a short time previously. They would also have some bearing on the cases where two or more techniques are used in the same sales interview. For instance, if both long and short circuit procedures were to be employed, their order might make some difference in the results.

Many limitations of the present experiment were discovered and would need to be taken into account in further research. The experimenter's rating scale for example, would come in for some revision. The item "questions" as originally designed apparently involved two factors—the presence or absence of questions and their quality. Similarly the rating of conversation might well be subdivided on the basis of frequency and the type of conversation involved. The experimenter was conscious of the difficulty in making judgments while actually delivering a sales talk. A possible improvement would be to have a third party watching the sale perhaps through a one-way screen and recording such things as posture and conversation on the part of the subject. However, from a practical standpoint, in the actual interview, the salesman has to make his own judgments.

Further extension of the method would naturally involve rotating the products and the techniques. All eight of the techniques, for instance, might be used for selling coffee. It is possible that a technique which would be well adapted for one product would not be so satisfactory for another.

The results of the experiment are subject to the limitations mentioned at the outset, but the method appears fruitful. The graphic rating scale is reasonably well adapted for the use of prospect and salesman in an experimental sales interview. The results are sufficiently significant to warrant the extension of similar techniques to the investigation of other factors in personal selling.

SUMMARY

Standardized sales interviews were conducted in the laboratory and both the subject (prospect) and the experimenter (salesman) filled out graphic rating scales at the conclusion of each interview. The subject's scales involved interest, feeling, confidence, acceptance of the proposition, salesman's attitude and persuasiveness of the interview. The experimenter's ratings included interest, conversation on the part of the subject,

questions asked by the subject, inspection of the product (if presented), and posture of the subject. The experiment was limited to four pairs of sales talks and comparison made only between the two members of each pair. Both members of a pair involved the same product and were equated so far as possible in other respects.

A demonstration proves superior to an oral elaboration in almost every respect and the differences in ratings are statistically significant. When a factual approach is compared with a short circuit appeal, the subject's ratings consistently favor the former. The experimenter's ratings are equivocal except that they indicate more conversation during the factual presentation. It is possible that rationalization leads the subject to rate the factual appeal more highly because it reflects favorably upon him. Comparing a breezy with a dignified approach, the average results are more equivocal and presumably reflect the fact that some subjects react favorably to a breezy approach and others unfavorably. When a domineering is compared with a friendly approach, the results consistently favor the latter by large and significant differences in the case of the subject's ratings.

Investigation was made of certain correlations between the ratings. The experimenter's estimate of the subject's interest shows no appreciable correlation with the subject's estimate of his own interest. Neither do any of the other variables which the experimenter rated correlate appreciably with the subject's expressed interest. If this result is corroborated by other experiments, it means that a salesman has difficulty in judging a prospect's interest by his observable behavior. Inter-correlations between the estimates made by the subjects are high except when the item "salesman's attitude" is involved. The order in which the products are presented apparently has some influence on the results but conclusions are not justified. The effectiveness of a given technique for a product may be influenced by previous experience with a different technique for that same product.

The experiment is presented primarily as a technique and is subject to the limitations that only one experimenter was employed and that there is a dissimilarity between laboratory and practical situations. It could be expanded, however, in both these respects. As far as it went, the results proved interesting and many of them were statistically significant within the present sample of data. This approach to personal selling appears worthy of further extension.

APPRENTICES' ATTITUDES TOWARD THEIR TRAINING AND THE CONSTRUCTION OF A DIAGNOSTIC SCALE[1]

H. E. GEIGER, H. H. REMMERS, AND R. J. GREENLY
Purdue University

INTRODUCTION

THE importance of the industrial worker's satisfaction with his job has become more and more apparent; the need for instruments to measure and analyze that satisfaction is well recognized. One approach to the problem is through the use of attitude scales.

A DIAGNOSTIC ATTITUDE SCALE

Thurstone, using the concept of the equally often observed difference, developed a technique for constructing scales designed to measure specific attitudes. Remmers modified the Thurstone technique by constructing a generalized scale to measure any one of a group of attitudes. This has been accomplished by the construction of generalized master scales which may be used to measure the same types of attitudes (3) and which yield points on a continuum assumed to represent the attitude measured. An attitude measured in this way is considered as a single attitude; although it is well understood that it is greatly affected by sub-, related, or auxiliary attitudes. In this article we describe the development and use of a scale which uses these auxiliary attitudes to make the scale a diagnostic instrument.

The single score obtained from a generalized attitude scale

[1] This investigation has been aided by a grant from the Josiah Macy Jr. Foundation and is to be published in detail in a bulletin of Purdue University. The present paper is being published by permission of Purdue University.

indicates only the intensity of favorableness or unfavorableness of the attitude. If, however, the instrument can be made to show also the type of statement most frequently agreed with, the factors influencing the attitude toward favorableness or unfavorableness may be suggested. If, in a scale designed to measure apprentices' attitudes toward their training, it can be shown that items relating to foremen are checked more often on the favorable end of the scale while items relating to instruction are checked more often on the unfavorable end, a basis for diagnosis is available.

An attitude scale which is internally diagnostic may be constructed with very little more work than that required for those already in use. Statements which are generalized in regard to a major attitude, but separately related to each of several subsidiary aspects must be edited. A few more statements than in the previous scales, perhaps twice as many, may be necessary in order to obtain statements which will represent a skeletal continuum of the whole general scale. Procedures for weighting the statements, computing values, and scale assembling may be handled in the same manner as before. Scoring may be done fully as easily as in the ordinary scale, the subsidiary attitudes being readily shown by a simple scoring design.

In the construction of a scale to measure apprentices' attitudes toward their training, the general attitude dealing with the whole program was fairly easily defined as the apprentice's attitude toward his training, but the more specific subsidiary attitudes were not so easily handled. However, after some consideration, the subsidiary attitudes were grouped under seven major headings, each dealing with a particular one of those aspects which were judged to be the more important in the training program. It must be kept in mind that these subsidiary divisions, while critically chosen, were nevertheless selected upon a more or less arbitrary basis. Under these headings were included attitudes toward the job proper, toward fellow employees, toward foremen or inspectors, toward

the promotional program, toward related technical instruction, toward general management, and toward collective bargaining.

CONSTRUCTION OF THE DIAGNOSTIC SCALE

Editing the Statements

One of the most important steps in scale construction is that of securing an adequate number of carefully selected statements. Such statements when scaled, must represent values extending from one extreme of a continuum of favorableness through the neutral point to the other extreme. Thurstone suggests the "attitude variable" which must be so stated that it may be spoken of in terms of "more or less" (5). The generalized scales have been constructed in such a way that, even though different attitude objects were measured, the "variable" remained consistent throughout the scale. In the construction of the present scale, the author had in mind not only a general attitude variable which would be consistent throughout the scale, but also sub-, auxiliary, or intra-attitude variables, corresponding to the subsidiary attitudes which were being considered. This made it necessary not only to select statements with values which would run from one end of the continuum to the other and be consistent with a general attitude variable, but, in addition, to select those statements in such a way that some would pertain to each of the intra-attitudes.

The usual number of preliminary statements selected has been about two hundred. It will readily be seen from the preceding paragraphs that a much larger number must be selected in order to be sure of having intra-scale statements for each of the subsidiary attitudes with scale values representing a sort of skeletal continuum of equally spaced items running throughout the scale, and, at the same time, of having the items of each intra-scale intermixed such that the items of the major scale may be equally spaced.

About five hundred statements were edited; however, after carefully ruling out those which seemed less desirable, four

hundred fifty-one statements were retained in the preliminary list. This list was then submitted in duplicate copies to five judges. This time the number was reduced to four hundred three. Scale values were secured from the ratings of college students, and, after careful criteria had been set up to eliminate undesirable ones, there remained 266 statements.

From the 266 statements were selected the statements for the final general scale which contains fifty-seven items with scale values ranging from 10.6 to 1.2 on an eleven point continuum. There are two equivalent forms, form A and form B, which were prepared for determining reliability. Each of the two forms is made up of six intra-scales, each of which pertains to a particular aspect of the working situation. Each of the intra-scales as well as the major scale has its neutral point near 6.0, statements on one side of the neutral point for any scale being balanced by statements at approximately the same distance from the neutral point and on the opposite side.

SCORING DESIGN

The six intra-scales are indicated by the numbers in the six squares, along with spaces for recording scores, at the top of the sheet (Fig. 1). Directly below each of the intra-scale numbers are squares opposite the statements which correspond to that particular intra-scale. From these squares, scores for the intra-scales may be obtained. Scores may be read directly from a key (scale values) held along the squares of any intra-scale. The score for the general scale may be obtained by computing the score for all of the items, regardless of intra-scale identities.

SCALE APPLICATION AND APPRENTICES' ATTITUDES

The final test of any instrument is the use to which it may be put. The scale, forms A and B of "A Diagnostic Scale for Measuring Apprentices' Attitudes," was applied to six groups of apprentices in as many industrial plants, as follows:

All of the groups except group five, which consisted of nurses

A DIAGNOSTIC SCALE FOR MEASURING APPRENTICES' ATTITUDES

H. E. Geiger

Edited by H. H. Remmers

Form A

Please fill in the blanks below.

Name ———————————— Sex M F (encircle one)

1	2	3	4	5	6

Age—— How long at this job? Years—— Months——
Directions: The following is a list of statements about your job and training program. Put an (X) in the square opposite any statement with which you agree, thus: ☒. Mark only those statements with which you agree. Your score will in no way affect your job.

☐ 1. This is an excellent job.

☐ 2. The foreman is liked extremely well by the men.

☐ 3. I am very happy with the responsibility in my job.

☐ 4. I hope that this company will continue to make good.

☐ 5. I like the inspectors very much.

☐ 6. Responsibility in my job gives me a sense of satisfaction.

FIG. 1.

in training, were general factory groups. The average age for all groups was 20.7 years.

GROUP	NUMBER OF APPRENTICES
1	83
2	49
3	32
4	5
5	23
6	35

RELIABILITIES

Reliabilities determined by the Pearson Product Moment Method were obtained for the general scale and for each of the intra-scales by applying forms A and B consecutively. It will be seen from Table 1 that the correlations between the two forms, A and B, range from .24 ± .07 for intra-scale II which

TABLE 1
Reliabilities—Forms A and B

	INTRA-SCALE	NO. ITEMS	N	MEAN$_A$	S.D.$_A$	MEAN$_B$	S.D.$_B$	r	P.E.$_r$
I	Job Proper	19	222	9.1	.53	9.2	.67	.49	.03
II	Fellow Employees	2	83	8.4	1.11	8.1	1.53	.24	.07
III	Foremen	15	223	8.5	1.37	8.7	1.49	.72	.02
IV	Promotion	7	187	8.1	1.65	8.1	1.79	.58	.03
V	Related Instruction	7	190	7.7	1.15	7.9	1.16	.45	.04
VI	Management	7	206	8.2	1.48	8.2	1.70	.31	.04
	General Scale	57	223	8.8	.58	8.9	.80	.60	.03

contains only two items, to .72 ± .02 for intra-scale III which contains fifteen items. The reliability for the general scale is .60 ± .03. The reliabilities for the intra-scales are fairly low. The reliability for the general scale is low for individual comparison but quite adequate for group comparisons. It should be pointed out that the groups measured were obtained from a rather select population of young people starting out on their vocational careers. This factor would tend to lower the reliabilities which greatly depend upon the variability of the group. Low standard deviations seem to substantiate such an assumption. No statistically significant differences were found between any two of the scales when the means of the distributions of the two forms were compared.

RELATION OF INTRA-SCALES

A second thing to be considered is the matter of the relationships between the various intra-scales. If any two of the scales were found to have a high correlation, it might be assumed that they measured the same thing and therefore one of them could be eliminated. It will be seen from Table 2 that the intercorrelations of the intra-scales are low. These correlations, determined by the Pearson Product Moment Method, range from .50 ± .05 to .0 ± .06; however, most of them range near the twenties, with an average of .21. Corrected for attenuation the correlations, except between I and II, range between .70 ± .10 and .0 ± .05, with an average of .39. Cor-

TABLE 2
Relation between Intra-scales

SCALE		II	III	IV	V	VI
I	N	133	226	213	211	225
	r	.50 ± .05	.14 ± .04	.38 ± .04	.26 ± .04	.31 ± .04
	r_∞	1.45 ± .39	.24 ± .08	.71 ± .07	.55 ± .09	.79 ± .10
II	N		132	127	128	132
	r		.0 ± .06	.11 ± .06	.16 ± .06	.13 ± .06
	r_∞		.0 ± .06	.29 ± .15	.47 ± .17	.48 ± .21
III	N			213	211	225
	r			.22 ± .04	.0 ± .05	.24 ± .05
	r_∞			.34 ± .07	.0	.51 ± .09
IV	N				200	212
	r				.21 ± .05	.34 ± .04
	r_∞				.41 ± .08	.50 ± .06
V	N					212
	r					.08 ± .06
	r_∞					.21 ± .12

relation between I and II when corrected for attenuation is 1.45 ± .39. It should be noted that intra-scale II contains only two items and has a very low reliability. It may be concluded that the intra-scales measure quite different aspects of the general attitude toward the training program. It must be noted, however, as mentioned above, that the group which constituted the population measured was relatively homogeneous in nature and not very large. These correlations, as well as those for reliability, would be raised by a more heterogeneous group.

That the intra-scales measure different aspects of the attitude toward the working situation is further evidenced by the significant differences found in comparing the various means of the distributions of the intra-scales. It is of importance here to point out the significant differences found between intra-scale I, relating to the job proper, and any of the other intra-scales. The job is always preferred, and in some cases, as in comparison with related technical instruction, it is decidedly preferred. This is consistent with Hoppock's conclusions

(2) and doubtless is of considerable consequence in the matter of scale validity. It seems reasonable that young apprentices who have been more or less carefully selected would be favorably inclined toward their chosen vocation. It should be noted also that all of the intra-scales deviate significantly from the general scale. These mean differences support the contention for diagnostic values of the scale.

ANALYSIS OF GROUPS

The scores from each group were analyzed individually and compared with the sum of the scores of all other groups thrown together as a norm. Because of lack of space only group 1 will be given in this article. Complete details and analyses may be obtained from the published study in the Bulletin of Purdue University.

While some of the groups showed significant differences both above and below the sum of all other groups taken as a norm, group one did not deviate significantly in any of the intra- or general scales (see Table 3). Between the intra-scales within group one, however, several differences were found. The score for intra-scale I, relating to the job proper, was significantly higher than any of the other intra-scale scores. The foremen were liked better than the related technical instruction. Related technical instruction in this group rates lower than any of the other aspects measured, and it is significantly lower when compared with the attitude toward the job proper or the attitude toward the foremen.

SUMMARY AND CONCLUSIONS

It was desired to construct a generalized scale for measuring attitudes of apprentices in industry. The possibility of organizing such a scale about several major subsidiary attitudes in order to make it diagnostic in nature was carefully worked out and the scale developed. Items for the scale were selected from four hundred and three statements which had been classified according to six different aspects of a training

TABLE 3

Intra-scales and general scale

Group 1 Compared with Total of All Others as a Norm

Form A

Comparisons of Mean Scores

	SCALES	$MEAN_1$	$S.D._1$	$MEAN_T$	$S.D._T$	DIFF.
I	Job Proper	9.1 ± .06	.54	9.4 ± .04	.47	.3 ± .07
II	Fellow Employees	8.3 ± .19	1.39	8.5 ± .10	.90	.2 ± .22
III	Foremen	8.5 ± .15	1.41	8.5 ± .11	1.35	0 ± .17
IV	Promotion	8.3 ± .16	1.38	8.0 ± .15	1.79	.3 ± .22
V	Related Instruction	7.7 ± .14	1.21	7.8 ± .10	1.12	.1 ± .15
VI	Management	8.2 ± .15	1.36	8.2 ± .13	1.55	0 ± .21
	General	8.8 ± .06	.53	8.8 ± .05	.62	0 ± .08

Differences between Scales

	II	III	IV	V	VI	G
	←	←	←	←	←	←
I	4.00* = 100†	3.75* = 100†	4.71* = 100†	9.33* = 100†	5.67* = 100†	3.75* = 100†
		←	←	←	←	
II		↑ .83 = 80	0.0 = 50	2.50 = 99	.42 = 67	↑ 2.50 = 99
			←	←	←	
III			.99 = 84	4.00 = 100	1.43 = 92	↑ 1.88 = 97
				←	←	
IV				2.86 = 100	.18 = 57	↑ 2.95 = 100
V					↑ 2.50 = 99	↑ 7.33 = 100
VI						↑ 3.75 = 100

* Indicates the critical ratio.

† Indicates chances in 100 of a true difference.

Arrow indicates higher mean.

program and which had been weighted by a large number of individuals. These items were balanced in various ways between two equivalent forms and within each of the forms. Items were set up on the sheet in such a way that scores could be easily obtained for each of the subsidiary groups of statements or intra-scales as well as for the total general scale.

The scale has been administered to six groups of apprentices and the results analyzed. The correlation between the two forms, A and B, for two hundred and twenty-seven apprentices was found to be $.60 \pm .03$; but, since the group was rather homogeneous, this correlation is probably much lower than would otherwise be expected. The intercorrelation between the intra-scales when corrected for attenuation is low, averaging .39, and seems to indicate that they really measure different aspects of the attitude toward the training program.

A considerable number of significant differences was found between means of the distributions of scores, both between groups and within each group.

BIBLIOGRAPHY

1. Henig, Max S. "Age at Entrance and Term of Stay," *Industrial Arts and Vocational Education*, Vol. 25, July 1936, pp. 203–205.
2. Hoppock, Robert. *Job Satisfaction*, Harper and Brothers, New York, 1935.
3. Remmers, H. H. "Studies in Attitudes," *Studies in Higher Education XXVI, Bulletin of Purdue University*, Vol. 35, December, 1934.
4. Sigerfoos, Charles C. "The Validation and Application of a Scale of Attitude toward Vocations," *Studies in Higher Education XXXI, Bulletin of Purdue University*, Vol. 37, December, 1936, pp. 177–191.
5. Thurstone, L. L., and Chave, E. J. "Measurement of Attitude toward the Church," *University of Chicago Press*, Chicago, Illinois, 1929, pp. 2.

BACKGROUND AND PERSONAL DATA AS FACTORS IN THE PREDICTION OF SCHOLASTIC SUCCESS IN COLLEGE*

CARRIE MAE SCOTT
Crowley, Colorado

THE aim of this study was to evaluate background and personal data items to determine their efficiency in the guidance of freshmen students entering Colorado State College of Education.

PROCEDURES

A questionnaire containing items relative to background and personal conditions was administered to the 280 freshmen students who attended Colorado State College of Education during the academic year 1934–1935. The students had completed a full year's scholastic work, and had taken all of the tests contained in the Entrance and Classification Examination for Teachers Colleges.

The intellectual status of each student was determined upon the bases of the American Council Psychological Examination and the combined raw scores on the English and Elementary Tests from the Entrance and Classification Examination for Teachers Colleges.

The permanent record cards filed in the registrar's office were used to find the average scholastic grade for each student in terms of a numerical value.

The various sections of the items in the questionnaire were

* Scott, Carrie Mae, *Background and Personal Data as Factors in The Prediction of Scholastic Success in College.* Unpublished Master of Arts Thesis, Colorado State College of Education, Greeley, Colorado, August, 1936. (The thesis was produced in collaboration with Dr. J. D. Heilman, Director of Personnel.)

Item	*Answer*	*Score*	
1. My father's principal occupation belongs to the following group:	(a) Professional	6	
	(b) Business	6	
	(c) Skilled	5	
	(d) Semi-skilled	5	
	(e) Unskilled manual	4	
	(No answer)	4	
2. The grade or grade group last attended by my father was:	(a) Grades 0–7	5	
	(b) Grade 8	5	
	(c) Grades 9–11	5	
	(d) Grade 12	6	
	(e) Years in college, 1–3	6	
	(f) Years in college, 4 or more	7	
	(No answer)	4	
3. My home is in:	(a) A city of 5,000 or more	6	
	(b) The country or a town of less than 5,000	5	
	(No answer)	4	
		B.	*L.*
4. The ONE school subject which I like best is:	(a) Art	4	9
	(b) Biology	5	7
	(c) Chemistry	7	6
	(d) Commercial	4	7
	(e) English and literature	6	4
5. The ONE school subject which I like least is:	(f) Foreign language	6	5
	(g) General science	5	6
	(h) History and civics	5	5
	(i) Home economics	3	6
	(j) Industrial arts	3	6
	(k) Mathematics	5	5
	(l) Music	5	7
	(m) Physics	6	6
	(n) Sociology	6	5
	(No answer)	3	4
6. I like the high school subjects in general:	(a) Very well	6	
	(b) Well	5	
	(c) Neither like nor dislike	4	
	(d) Dislike	2	
	(No answer)	2	
7. The number of hours per week which I devoted to the preparation of my lessons during the last year in high school was:	(a) 0–4 hours	4	
	(b) 5–24 hours	5	
	(c) 25 or more hours	5	
	(No answer)	3	
8. My age at high school graduation was:	(a) 15 or under	7	
	(b) 16–17	6	
	(c) 18	6	
	(d) 19	4	
	(e) 20 or over	3	
	(No answer)	2	

Item	*Answer*	*Score*
9. The number of years which elapsed between my high school graduation and college entrance was:	(a) Less than one	5
	(b) One or more	4
	(No answer)	4
10. The degree of interest shown by my parents in my high school work was:	(a) High	6
	(b) Average	4
	(c) Low	4
	(No answer)	3
11. My admiration of the capable student is:	(a) Very high	5
	(b) High	5
	(c) Average	3
	(d) Low	3
	(No answer)	2
12. My admiration of the industrious student is:	(a) Very high	5
	(b) High	5
	(c) Average	3
	(d) Low	3
	(No answer)	2
13. The ONE type of problem on which I most felt the need of advice during my high school course was:	(a) Educational	4
	(b) Family relations	5
	(c) Financial	5
	(d) Health	5
	(e) Moral	5
	(f) Religious	6
	(g) Sex	6
	(h) Vocational	6
	(No answer)	5
14. That ONE of the following traits which was most helpful to me for successful work in high school was:	(a) Native ability	7
	(b) Industry	6
	(c) Promptness, regularity	6
	(d) Character, personality	6
	(e) Ambition	6
	(f) Sincerity, honesty	6
	(g) Interest	6
	(h) Interest in reading	5
	(i) Study habits	5
	(j) Friendliness	5
	(k) Initiative	5
	(l) Willingness, cheerfulness	5
	(m) Health	4
	(No answer)	4
15. That ONE of the following traits which handicapped me most in doing successful high school work was:	(a) Impatience	8
	(b) Procrastination	8
	(c) Lack of effort	8
	(d) Overconfidence	7
	(e) Lack of initiative	7
	(f) Shyness	7
	(g) Lack of self-confidence	7

Item	*Answer*	*Score*
	(h) Outside interests	6
	(i) Slowness, carelessness	6
	(j) Lack of interest	6
	(k) Deportment	6
	(l) Poor study habits	6
	(m) Stubbornness	6
	(n) Lack of natural ability	6
	(o) Dislike of teacher, subjects	5
	(p) Health, physical defects	5
	(q) Poor in some subject	4
	(r) Distance from school	2
	(s) Finance	2
	(No answer)	4
16. I decided upon my life work at the age of:	(a) 15 or under	6
	(b) 16–17	5
	(c) 18–19	4
	(d) 20–21	2
	(e) 22 or over	4
	(No answer)	2
17. I am planning to attend college:	(a) One year	4
	(b) Two years	5
	(c) Three years	6
	(d) Four or more years	6
	(No answer)	3
18. Financial provision for my attendance at college has been made for:	(a) Two years or less	5
	(b) Three years or more	6
	(No answer)	4
19. My ONE most important reason for going to college is:	(a) Engage in college athletics	2
	(b) Enjoy social life at college	3
	(c) Satisfy the desire of parents, relatives or other persons	3
	(d) Improve my social standing in life	4
	(e) Improve my opportunity for earning a living	5
	(f) Fulfill my desire for learning	6
	(g) Prepare for some worthwhile life activity	7
	(No answer)	4
20. The grade level on which I am planning to do most of my teaching is:	(a) Below the fourth grade	4
	(b) Grades 4 to 8	5
	(c) The high school	6
	(d) Not planning to teach	5
	(No answer)	4

assigned weighted numerical values based upon percentile scores obtained on the combined English and Elementary Tests.

A copy of the questionnaire together with the assigned weight values follows:

A total background and personal data score was computed for each student. The scores were correlated with average scholastic grades, American Council Psychological Examination scores, and the combined English and Elementary Test scores using the Pearsonian formula. The intercorrelations among the several factors together with the arithmetic means and sigmas of each factor are illustrated in Table I.

TABLE I

Intercorrelations Among Factors for 280 Freshmen of Colorado State College of Education Together with the Arithmetic Means and Sigmas of Each Factor and the Probable Errors of the Correlations

VARIABLES	AVERAGE SCHOLASTIC GRADES	QUESTIONNAIRE SCORES	AM. COUNCIL PSYCHOLOGICAL EXAM. SCORES	ENGLISH AND ELEMENTARY TEST SCORES
1	2	3	4	5
Questionnaire scores	.370			
P.E.$_r$	.034			
Psychological Exam. scores	.637	.330		
P.E.$_r$	.024	.039		
English and Elem. Test scores	.724	.395	.925	
P.E.$_r$	.019	.034	.006	
Arithmetic means	3.12	97.82	165.4	459.40
Sigmas	.626	4.82	49.95	24.40

The correlation obtained between the questionnaire scores and grades was .370 with a P.E. of .034; between questionnaire scores and American Council Psychological Examination scores .330 with a P.E. of .039; and between questionnaire scores and combined English and Elementary Test scores .395 with a P.E. of .034. The correlations indicated that the background and personal data items measured by the questionnaire were of some use for the prediction of probable success in college, both in passing the entrance examinations and in classroom achievement.

A regression equation using the variables; average scholastic grades, Psychological Examination scores, and the combined English and Elementary Test scores, formed the basis for a prediction of probable college grades. The multiple correlation was .723 with a sigma estimate of .432 and a P.E. estimate of .291.

The regression equation in score form became:

$$X_1 = .003X_2 + .005X_3 + 38$$

The probable grades of fifty students were predicted on the basis of the regression equation and compared with actual grades. The median difference between the predicted grades and the actual grades was .504, or about one-half of a letter grade.

A four variable regression equation was computed for the same group adding the questionnaire scores. The coefficient of multiple correlation was .743, an increase of .020. The sigma of the estimate was .419 and the P.E. estimate .282.

The regression equation in score form was:

$$X_1 = .013X_2 + .003X_3 + .005X_4 - 90$$

Scholastic grades were predicted for the same fifty students upon the basis of the four variable equation and compared with actual grades. The median difference between predicted grades and actual grades was .503. This was an increase of .001 in accuracy of prediction.

The significance of each section of the items in the questionnaire was determined by comparing it with average scholastic grades. The sections of each item were grouped with relation to their logical groupings and with reference to average grades. The median, Q_3, Q_1, and Q were found for each distribution. Differences with their P.E.'s were found between the medians of the various groups.

The groups who made the highest and lowest median grades are illustrated in Table II.

Many of the items in the questionnaire seemed to possess a value for the determination of scholastic success in college. These items concerned: the size of the home population unit; best and least liked high-school subjects; liking for the high-

TABLE II

Median Scholastic Grades of Freshmen of Colorado State College of Education Who Answered the Items of the Background and Personal Data Questionnaire in a Specified Manner

ITEM	NO.	HIGHEST GRADE GROUP	MDN. GRADE	NO.	LOWEST GRADE GROUP	MDN. GRADE
1	2	3	4	5	6	7
1. Occupation of father	8	Unskilled	3.40	59	Business	2.97
2. Education of father	52	High school	3.17	20	College	2.85
3. Size of home population unit	116	City of 5000 or more	3.12	161	Small town or country	3.02
4. Best liked high school subject	9	Physics	3.08	10	Gen. Science	2.95
5. Least liked high school subject	3	Biology	3.60	9	Commercial	2.70
6. General liking for H. S. subject	114	Very well	3.35	20	Neither liked nor disliked	2.87
7. Hours preparation per wk. in H. S.	21	25 or more	3.18	225	5–24 hours	3.05
8. Age at high school graduation	4	15 or under	3.40	5	20 or over	2.50
9. Years between H. S. grad. and Col. Entr.	18	3 or more years	3.16	43	1 year	2.96
10. Degree of parent interest in H. S.	201	High	3.10	74	Average	2.96
11. Admiration of capable student	159	Very high	3.10	23	Average, low	2.95
12. Admiration of industrious student	210	Very high	3.08	8	Average, low	2.90
13. Types problems advice needed in high school	7	Religious	3.50	14	Moral	2.93
16. Age decision to teach	85	15 or under	3.14	27	18–19	2.97
17. Time planned to attend college	193	4 or more years	3.10	17	3 years	2.97
18. Years of financial provision	212	2 or less	3.08	47	3 or more years	3.05
19. Reasons for going to college	4	Prepare for worthwhile life activity	3.40	15	Improve social standing in life	2.97

school subjects in general; age at high-school graduation; admiration of the capable student; admiration of the industrious student; years planned to attend college; and, reasons for going to college. A difference that was large enough to be statistically significant was found in the item concerning the size of the home population unit.

CONCLUSIONS

The study purported to evaluate the efficiency of a background and personal data questionnaire in the prediction of freshman success in college.

The sections of each item were assigned weight values. A coefficient of correlation of .370 was obtained between total questionnaire scores and average scholastic grades; .330 between questionnaire scores and American Council Psychological Examination scores; and .395 between questionnaire scores and combined English and Elementary Test scores. The intercorrelations indicated that the questionnaire was of some value for predicting scholastic success in college.

A regression equation based upon the Entrance Examination scores was used to predict the probable grades of fifty students. With questionnaire scores as an added variable, grades were predicted for the same students. A higher degree of accuracy in prediction resulted when the questionnaire scores were added to the Entrance Examination scores.

The answers to the items of the background and personal data questionnaire were grouped with relation to their sections and to average scholastic grades. Difference in medians were found with their P.E.'s. One item contained a reliable difference in medians. Other items contained differences great enough to indicate that they might have a bearing upon success in college. More items would probably have been found significant with the use of a larger group.

The background and personal data questionnaire had a bearing upon freshman achievement in college. However the predictive value was not great enough to justify its use for that purpose.

THE TRUTH ABOUT THE LIE DETECTOR

CHRISTIAN A. RUCKMICK
University of Iowa

TRULY it is presumptuous even for a psychologist to try to attract a group of readers composed, for the most part, of his own colleagues with the title that I have chosen for this paper. I might as well abjectly confess, however, that as an unfertilized ovic philosopher, in undergraduate days, I produced a portly paper on the *Nature of Truth*. Be that as it may, most of us will concede that much untruth has been published in the press about so-called lie detectors of one sort or another and that a beginning should be made in the direction of a critical study of techniques commonly used for this purpose. Larson who has done much practical work with this technique in the actual environment of the court room, has led the way with an excellent monograph on the subject.[1] Others like Marston and Keeler have discussed their procedures and results in the popular magazines or in scientific journals. Perhaps most encouraging in this regard is Larson's statement in the preface of his book:

"Even if the reader may be stimulated to experiment with the technique because of disbelief in the procedure discussed, and can ultimately show experimentally that such procedure is impractical, such a discovery and research will have been well worth while."[2]

Elsewhere he carefully surveys the possibilities of the technique which he uses, referring preferably to the 'cardio-pneumo-psychograph' rather than to the 'lie detector.' At present he is opposed "to the introduction of deception-test

[1] J. A. Larson, *Lying and Its Detection*, A Study of Deception and Deception Tests, Univ. of Chicago Press, 1932.

[2] *Ibid.*, p. xiv.

records into court procedure as the technique is not infallible and may never be accurate enough to warrant courtroom use. There can never be in this technique such accuracy as is possible in the case of fingerprints, and so long as error is possible, these records should not be introduced for wrangling sessions by opposing attorneys, as they would only further mystify the jurors and might lead to a flagrant miscarriage of justice."[3]

When in the hands of competent persons he claims that the technique compares "favorably with other police or legal methods of investigation" in the preliminary "interrogation of suspects and their selection." With an admitted success of over 90 per cent accuracy in some ten thousand suspected cases, mostly in terms of a confession, however, the claims on behalf of this technique become impressive.[4]

The history of the problem of applying psychological principles to the detection of crime is almost as old as the present century. Popular attention was first called to the desirability of utilizing experimental procedures for this purpose by Hugo Münsterberg in 1907 and 1908 when he wrote a series of magazine articles on the subject which were later published under the famous title, *On the Witness Stand.*[5] Here we find briefly described some preliminary experiments on word associations, on testimony, and on emotional disturbances in breathing, circulation and involuntary movements. Even the galvanometer is mentioned in this connection.[6] But the question of lying as a special phase of the psychology of testimony probably became the center of interest when an Italian psychologist, Benussi, published in 1914 a study of the changes in respiration produced by lying.[7] Previously several investigations had been focused on the effect of mental processes and conditions on

[3] *Ibid.*, p. 415.

[4] L. Keeler, How science solves crime: the polygraph, *Hygeia*, 10, 1932, p. 740.

[5] H. Münsterberg, On the Witness Stand, N. Y., 1925.

[6] *Ibid.*, p. 129–130.

[7] V. Benussi, Die Atmungssymptome der Lüge, *Arch. f. d. ges. Psychol.*, 31, 1914, pp. 244–273.

bodily responses. Benussi found that the ratio of inspiration to expiration was generally greater *before* truth-telling than that *before* lying, with an inverse relationship holding for each case *after* the verbal statement had been made. In 1917 Marston suggested the use of records showing changes in systolic blood pressure[8] and Burtt followed in 1921 with some experiments tending to disprove Benussi's results and lending support to the belief that blood-pressure offered in some cases a better diagnostic technique.[9] Lately, with the perfection of galvanometric apparatus and techniques, this approach is again showing increased interest. There is no time and no occasion here to review critically the sundry techniques and data thus far obtained. This has been done at some length in my recent book.

It has occurred to me, however, that controlled experimentation on some particular phases of the problem of lie detection is in order. For example, in legal procedure, it is not always clearly established what the truth is or what is the base-line of reference in terms of which lying or falsification of fact occurs. As Larson admits, even when a confession is forthcoming as the result of a preliminary investigation by the psychologist, that confession, upon advice of counsel, may be repudiated. On the other hand, a confession thus obtained and not repudiated may enmesh an innocent person whose mind may have become so befuddled and bewildered that he can easily believe himself to be guilty of the charge. In a laboratory situation the truth is an easily ascertainable fact. On the negative side it must be admitted that naturally the conditions of the courtroom can not be reproduced *in toto*. Even with the utmost desire on the part of the experimenter to make the situation as grave and as potentially emotional as the corresponding situation is in everyday life, it is well nigh impossible to duplicate such an eventful occurrence. We must, however, bear these deficiencies in mind and proceed as analytically and guardedly as possible, so that

[8] W. M. Marston, Systolic blood-pressure symptoms of deception, *J. Exper. Psychol.*, 2, 1917, pp. 117–163.

[9] H. E. Burtt, The inspiration-expiration ratio during truth and falsehood, *J. Exper. Psychol.*, 4, 1921, 1–23.

we may know what the limitations as well as the advantages of our instrument and tools are. To that end we have begun a series of investigations on the detection of lying with the galvanometric techniques.

These problems fall within a frame of similar researches which have been published in monograph form for about a decade. A new volume of such studies has recently appeared.[10] The galvanometer used was the newest model of the Greenwald apparatus.[11] The electrodes were of the dry sheet-copper type also already described in the literature.[12] There were altogether 89 subjects, most of whom were advanced undergraduate or graduate students in psychology. Some of them observed two or more times, giving us a total of 96 series of responses. There were about twice as many women subjects as men.

The procedure was gradually modified. In a preliminary series of trials numbers from 1 to 10 were printed on cards 3 inches square and the subject would be asked to select one card at random from the pack. Previous to this the instrument had, of course, been balanced and the electrodes affixed to the palm of the subject's left hand in the customary manner. When the subject's position on the galvanometer scale had been adjusted and had become fairly stable, the experimenter would say "is it —?" to which the subject had been instructed to say 'no' with as uniform inflection and intensity as possible and without shaking the head. The order of the numbers was haphazard and the experimenter's voice was also kept as uniform in inflection and intensity as possible. The result was that the subject lied when the number which he held was called. The deflections were recorded once for each item and as many

[10] C. A. Ruckmick (editor), Studies in General Psychology, Vol. II, Univ. of Iowa Stud. in Psychol., *Psychol. Monog.*, 48, 1936, No. 2, pp. 1–76.

[11] D. U. Greenwald, An apparatus for measuring the electrodermal response, *Amer. J. Psychol.*, 47, 1935, 682–685; Electrodes used in measuring electrodermal responses, *ibid.*, 48, 1936, 658–661.

[12] C. A. Ruckmick, A new electrode for the Hathaway galvanic reflex apparatus, *ibid.*, 42, 1930, 106–107.

as 3–5 times on numbers for which relatively large deflections were recorded.

It became increasingly clear that the implied ordinal sequence of numbers was a disturbing factor and that simple mathematical relations were also of this nature. When, for an example, 8 was being held and the experimenter said 7 or 9, the proximal relationship often was cognized, or if 4 were called out, the subject introspectively reported that he had noticed the factorial or other relationship.

We therefore discarded numbers and substituted simple three-letter monosyllabic words of fairly low affective value and sufficiently non-related for our purposes. They were: NOR, AND, CAN, FAR, WAS, TOP, ARE, LET, NOW, and HAD. To overcome the excitement of initially presented words, we added PUT and SAT which were not in the series from which the subject drew his card and which were always presented first in the series. To avoid the possible changes in the experimenter's pronunciations and inflections, a duplicate set of cards was prepared, $3'' \times 5''$ in size, on which the words were printed, one on each card, including the two initial words. Each card was presented in front of the subject after a 'ready' signal.

The underlying purpose of these experiments was to examine the procedure critically from the methodological point of view in much the same attitude as I examined the questionnaire procedure some years ago.[13] In other words, we were not primarily concerned with the 'efficiency' of the technique, nor more specifically with the problem of 'getting something on' the subject. To dispose of the matter of 'efficiency' first we may note that in the preliminary series in which the experimenter gained practice with the procedure, and in which there were certain material defects, we scored 66 per cent correct judgments. With the words we achieved a score of 78 per cent correct judgments. This score is increased to 83 per cent if we eliminate the work of one of the judges who was an under-

[13] C. A. Ruckmick, The uses and abuses of the questionnaire method, J. Applied Psychology, 19, 1930, pp. 32–41.

graduate student. The score is slightly raised above this if we make certain other eliminations which would not necessarily occur in everyday life. Among these are repeated observations from the same subject, usually when improvements had been made in the apparatus or technique or both. This sort of repeated observation is not always effective, however, in lowering the score of judgments.

Introspections further revealed the fact that some of the errors in judgment on the part of the experimenter are due to the fact that, in spite of preliminary care, certain words had internal similarities which were noticed by the subjects, *e.g.*, NOW and NOR, ARE and FAR. These produced disturbing deflections like those mentioned in regard to the preliminary series. Other disturbing factors which we detected were temporary emotional situations such as highly affectively colored ideas concerning course examinations, family problems and the like. It should be remembered that the affective life of an individual who is investigated by this technique is not always under the control of the experimenter. There are time and opportunity for the subject to wander off to other fields of interest to such an extent that the galvanometer will sometimes register deflections that are inexplicable except in terms of an honest and objective introspective report. These accounts were therefore extremely helpful in this regard, but, of course, the experimenter's judgment as to which card was retained was always declared before the analytical report of the experience was called for. In one case a subject, who was not only familiar with the literature but had participated in discussions concerning lie detectors, calmly proceeded intentionally to develop an excitement concerning a word which was in the series but not on the card retained. She could not thereby prevent a large deflection in connection with the word represented by the retained card, but she was able to cause a large deflection for the 'wrong' word, which misled the experimenter.

Accordingly we tried to discover whether this strategy would

be successful if we instructed the subject to become excited on some word other than the one represented on his card. This was tried in about a half dozen cases to date. This trick is generally successful in "throwing the experimenter off" until it is discovered that on repeated presentation of the words a difference begins to appear in the deflections. The word actually held usually causes a decrease in the deflection upon repeated presentation whereas the 'false' word is accompanied by increasingly large deflections.

Another marked though incidental result of these studies is the individual pattern of response with different subjects. Women generally give the larger deflections. Then there are great differences in the amount of deflection in individual cases. At one extreme is a girl who gives 0 deflections on the scale for all words except for the word she is holding which gives a reading of 25. At the other extreme is the person who gives a deflection of 25 or 30 for all words and 100 for the word withheld. It should be noted that a reading of one scale division is equivalent on this apparatus to a drop in resistance of approximately 300 ohms. A sweep from L to R across the entire scale accordingly represents a drop of 30,000 ohms. In a few instances a decision had to be made on the basis of an average difference of only 2 or 3 scale divisions. The problem of individual patterns of response, corresponding to brain-wave patterns, is being given further attention in a graduate thesis. There is every reason to believe tentatively that under a variety of conditions an individualized pattern is discernible.

This leads us to a final observation, namely that even under laboratory conditions the problem of lie detection is surrounded by a subjective factor of interpretation which is not yet reducible to quantitative terms. The question hinges on the increasing conviction, already expressed in my recent book,[14] that *patterns* of response are the criteria that are obviously used both in electrical manifestations and in the circulatory

[14] C. A. Ruckmick, *The Psychology of Feeling and Emotion*, N. Y., 1936, p. 342f, 371f.

and respiratory processes. As demonstrated in this work, the type of the deflection, what comes before it and what follows it, and the speed with which it occurs is as indicative as the amount of the deflection. It is as important here as it was in the diagnostic word-association test when we took reaction times not only to tell-tale words but to harmless words which followed. When a pattern of response is the objective datum, then it follows that a certain amount of practice effect is apparent in the process of judgment. This was noticeable among the three experimenters who independently worked as judges in this investigation: the number of correct judgments and the speed of judgment rapidly increased with practice after taking the various factors mentioned into account. When more detailed work of a critical nature becomes available, it ought to be possible to develop an equation in analytical geometry which will quantify the data obtained. At present the subjective factor of a perceptual order is still in the picture.

The data and conclusions reported are only a beginning. We are proceeding to investigate other phases of the problem, including lying in the affirmative, rather than in the negative, and also the detection of knowledge when there is no verbal response and consequently no lying in the usual sense—merely the withholding of information, when only the deflections will be available. The excitement of having lied will then be replaced by whatever excitement there is present when certain information is not externally revealed. While all subjects were told not to change the tone of their voices from one 'no' to the next and not to shake the head, it might be possible that small cues were picked up by this means by the experimenter. It is very doubtful from direct observation whether there were any such cues, but the possibility of such cues can be eliminated. There are many other angles of attack which are also significant for this experiment and which need further pursuit.

Altogether I feel that in the hands of an expert psychologist who knows the texture of the human mind this technique should prove valuable for use in cases of crime or misdemeanor. Ordi-

nary subjects are not likely to controvert the issue by getting themselves voluntarily excited over items that have no bearing on the event which is being investigated, as was the case with a few of our subjects. If they do, it may still be possible to detect the subterfuge in the manner that I have indicated. Constant allowance has to be made for sharp individual differences, however, and for the perceptual pattern in terms of which the diagnosis is made by the experimenter. It is noteworthy that one of our most skilful subjects, a Hindu who has made a long study of tension and relaxation with the galvanometer and has been able voluntarily to cause a deflection, was nevertheless 'detected' by our technique. Since, however, the instrument is not as reliable for purposes of identification as facial photography and finger printing, it becomes a dangerous weapon in the hands of any but the most competent persons. The situation is in the same category as are many other techniques including mental testing: only those who can see beyond the actual scores and interpret these scores in the frame of the individual mental life are competent to pass judgments.

DRIVER CLINICS IN THE FIELD

HARRY R. DeSILVA AND RALPH CHANNELL
Harvard Bureau for Street Traffic Research

DRIVER clinics consisting of a set of testing apparatus and a trained supervisor, under the sponsorship of the Harvard Traffic Bureau, have during the past year and a half operated in the following sixteen states: Massachusetts, Connecticut, New Hampshire, Vermont, Maine, Rhode Island, New York, New Jersey, Delaware, Virginia, West Virginia, North Carolina, South Carolina, Tennessee, Pennsylvania, and Alabama. California has its own clinic which we installed in September, 1936. The Wichita Police Department also has a permanent clinic operated by one of our trained men.

During the month's visit to a state the services of the clinic and supervisor have been furnished without charge.[1] The only demands on the sponsoring state department have been to supply space in which to operate, to furnish the assistance of examiners or troopers in testing, and to print a small educational pamphlet to distribute to persons attending the clinic.

PURPOSES OF A CLINIC

The purposes of our driver clinics are as follows:

1. To gather basic information from which to determine the most significant human causes of accidents. A data sheet is filled out by each driver covering such items as age, sex, mileage, driving experience, nationality, education and others. From these forms we have learned many things about the characteristics of accident repeaters as compared with unselected drivers.
2. To isolate the defects of accident repeater drivers by comparing their performances with those of unselected or accident free drivers.

[1] Through the courtesy of the Automotive Safety Foundation.

3. To give each driver personal advice on his defects and how to overcome or compensate for them. To show each violator or accident driver the connection between his defects and accidents. To supply each driver with an educational pamphlet containing an individual diagnosis of his weaknesses so that he may compare himself with other drivers of his acquaintance.
4. To encourage if possible the establishment of a file of all drivers tested so that their subsequent records may be followed up. Wherever such files have been kept, great improvement has been found in the subsequent accident and violator records of drivers attending the clinic.
5. To give new drivers a supplementary examination (in addition to the regular department, road, law, and eye test) with the best available apparatus. When requested, to inform and instruct examiners in the latest methods of psychological testing and educational interview technique.
6. To test drivers of state-owned cars, state and municipal traffic policemen and any other special groups. In one state a large group of applicants for the position of state patrolman were put through the clinic as a part of their general examination.

THE CLINICAL APPARATUS

The physical equipment used in a clinic consists of the following tests:

1. Vigilance Test
   a. Steering Test
   b. Braking Reaction Test
   c. Combined Braking and Steering
2. Universal Visual Test to Measure
   a. Acuity of eyes separately
   b. Depth Perception
   c. Astigmatism
   d. Color Vision
   e. Coordination
   f. Tunnel vision
   g. Glare Sensitivity
   h. Glare recovery
3. Speed Estimation Test
4. Hearing

Two of our latest model instruments are shown in the accompanying illustrations (Figs. 1 and 2).

FIG. 1.
The Vigilance Test.
To measure braking reaction time, steering and the ability to do several things at once.

THE CLINICAL TECHNIQUE

During the past three years we have developed a technique of diagnosis and treatment by personal interview which is an indispensable part of the operation of each clinic.

A driver clinic, like a medical clinic, is no more important than the trained specialist who supervises it. Our experience

has shown us that it is necessary to have clinical equipment supervised by a specially trained man. Each driver presents an individual problem. He must be encouraged to put forth his best efforts in every test. He must also be inveigled to talk about himself, particularly, to evaluate objectively his driving.

FIG. 2.
Universal Vision Test.
To test the most important visual factors involved in driving.

The individual scores from different tests may not be important separately, but the composite picture or profile of all these scores is of significance. The training and experience of a supervisor are indispensable in the diagnosis of past violations and accidents of the subject in terms of the composite clinical picture. But in addition to specialized training the clinic

supervisor must also have a pleasing personality in order to tender personal advice in an acceptable way.

PROCEDURE IN A STATE

The sponsoring agency in each state has generally been the motor vehicle department. When safety activities were in the hands of a supervisor of safety the clinic has operated under his direction. Where safety activities were under the highway patrol or police the clinic has cooperated with them.

Each driver clinic fits into a panel delivery truck[2] which conveys it from one location to another. The itinerary for a state is usually arranged in conference with the clinic supervisor. As many of the important cities of the state are covered as time will permit. In larger cities where the attendance at the clinic is heavy, the assistance of two or three examiners, or state or municipal policemen have been furnished.

The apparatus can be housed in a room approximately 15×15 feet. The clinic has been set up in space supplied by motor vehicle departments, state patrols, city police, fire departments, libraries, high schools, city buildings, universities and even vacant stores.

Various means have been used to call in people to the clinic. In a number of states the motor vehicle department has sent out letters to violators or accident repeaters requesting them to attend the clinic voluntarily for their own benefit. In three or four states all drivers called in for hearings by the motor vehicle department have been required to attend the clinic. Many volunteers have been attracted as the result of newspaper and radio publicity and from witnessing demonstrations at club meetings and schools. A considerable number of persons are also sent by people attending the clinic.

As far as possible the work of the clinic in each state is summarized and a report issued to the sponsoring authority. Some of the results of our driver clinics have been discussed in the articles mentioned in the bibliography.

[2] Vehicles have been generously supplied by the Ford Motor Company and the Dodge Brothers Co.

RESULTS

We may summarize the results of the driver clinic in motor vehicle departments as follows:

1. *To gather information about human causes of accidents:*

A. There are accident repeaters. A recent Bureau of Public Roads' study in the Connecticut Motor Vehicle Department has shown that over a period of six years 3.9 per cent of all drivers had 37 per cent of the accidents. These results prove that there is a small group, less than 5 per cent, of drivers who intermittently over a period of years pile up a surprising total of accidents.

B. Accident repeaters can be geographically located. A Detroit traffic survey shows that accident repeater drivers are not uniformly distributed. In two small sections of the city 51 per cent of the registered cars were in accidents while in other sections 5 per cent were in accidents.

C. Accident repeaters as a group can be improved. In two states where a check has been made on the records of a large number of drivers subsequent to attending our clinic, the results show a remarkable improvement in group performance. In one state a group of accident drivers had a seventh as many accidents as during a similar preceding period. In another state the drivers have had one twelfth as many accidents as they had during the previous year.

D. Accident repeaters as a group are less well educated. Our results show that accident repeater drivers have from two to three years less schooling than unselected drivers.

E. Foreign born drivers appear in larger proportion than expected among accident repeaters. This is to be explained on the basis of poorer education, less exposure to safety education, and more poorly maintained cars.

F. Accident incidence varies according to age. The peak of accidents in proportion to number of drivers is in the early twenties. The safest age for drivers is in the 40's and 50's. The accident incidence usually increases beyond 55. Pedestrian accidents are particularly high from 45 to 65.

G. Exposure in terms of extensive yearly mileage cannot explain accident repeaters. Accident repeaters as a group report approximately the same yearly mileage as unselected drivers. Among the repeaters are two groups: those with high mileage whose greater exposure leads to accidents, and those whose low mileage results in insufficient practice to keep in condition.

2. *To reeducate adults by personal contact:*

Our results in testing many thousands of drivers have shown that the only way to educate adults is to tackle them individually. Children can be coaxed or threatened into good behavior. Not so with adults. They must be treated as individuals especially when it comes to improving their behavior at the wheels of their cars.

The results of the previously mentioned Connecticut study also indicate that over a period of six years 80 per cent of all drivers had no reported accident. This means that the great majority of drivers have accidents over long periods of time and only an occasional violation. Naturally such prudent, fairly skillful drivers do not regard themselves as the subject of safety propaganda. Having a justifiable pride in their driving, they can be helped only by the personal interview approach based upon comparative measurements of skill under driving conditions rather than upon general rules of good driving.

Most adult drivers are fairly prudent, conscientious, proficient drivers. They need insight into:

1. Their lapses as affected by sensory or motor defects of which they are uninformed, or by their temporarily lowered vigilance resulting from distraction, intoxication, or fatigue.
2. Road and traffic conditions such as slippery roads, irresponsible pedestrians, etc.
3. Defects in their car from age, or poor maintenance.

Only by personal contact can individual drivers be encouraged

to analyze their driving, to overcome or compensate for any personal characteristics or external conditions predisposing them to accidents.

Let us cite some examples of human defects which have led to accidents:

An investigation by one of our driver clinics in the California Motor Vehicle Department showed that fully twenty

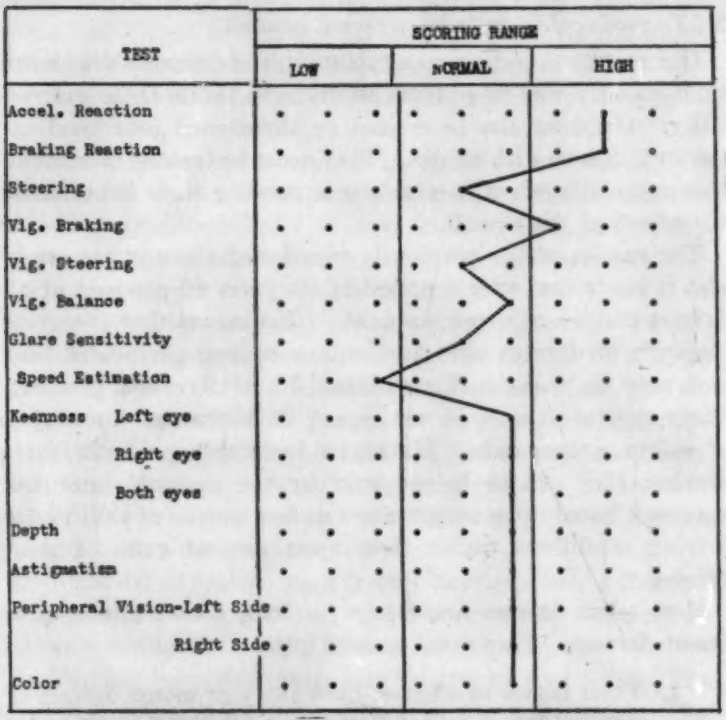

FIG. 3. Harvard Driver Clinic. Profile of Driver Skill.

Name _____ Date *June 16, 1937*

City and State *Providence, Rhode Island*

Age *51* Sex *Male* Examiner *R. Channell*

Seven rear end collisions to his car in one half year. Brake salesman. Delights in showing off what good brakes he has. Fast reaction time also.

per cent of all drivers in fatal accidents had poor vision in one eye. The seriousness of this condition was accentuated when it was found that in every case the accident occurred on the side of the weak eye. You can imagine the great value of this

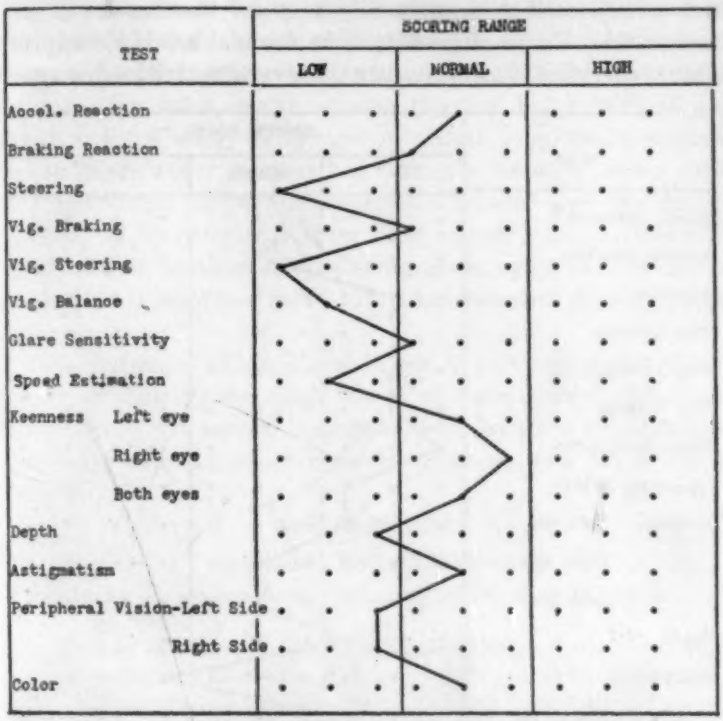

FIG. 4. Harvard Driver Clinic. Profile of Driver Skill.

Name _____ Date *July 9, 1937*

City and State *Burlington, Vermont*

Age *50* Sex *Male* Examiner *R. Channell*

Two reports of driving after he had been drinking, no convictions. Probably stopped because of his poor steering ability, which gave impression he was drunk. Has been driving three years about 4,000 miles per year. Born of foreign parentage. Eight years' education. Second test also poor.

diagnostic information to the clinical supervisor in his job of personally reeducating these drivers.

In Vermont a patrolman brought in an intoxicated driver he had picked up a few blocks from the clinic. The inebriated driver welcomed the opportunity to prove that he was in perfect fettle. To his discomfiture he found himself utterly unable to carry out simultaneously the several activities demanded

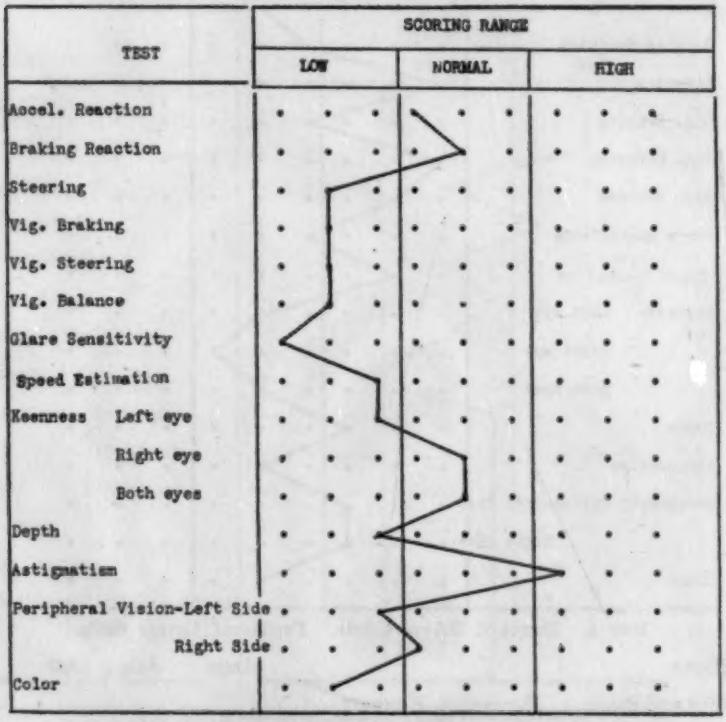

FIG. 5. Harvard Driver Clinic. Profile of Driver Skill.

Name _____ Date *May 4, 1937*

City and State *Providence, Rhode Island*

Age *65* Sex Male Examiner *R. Channell*

Five serious accidents in 11 years besides many minor ones.

by our vigilance test. After watching a normal person make a good score on the test he confessed his amazement at the demonstrated deterioration of driving ability from alcohol and vouchsafed that he would never again drive after drinking.

A case in Rhode Island (see Fig. 3) illustrates the need of compensating for superior ability along one line. A salesman came in bragging of his fast reaction time and efficient booster brakes. His score on the braking reaction test confirmed his statement. A study of the motor vehicle department records indicated, however, that he had been the victim of seven rear end collisions. After the supervisor pointed out the disadvantage, in his case, of a faster than normal reaction time and super brakes, he went away vowing never again to brag about or get into trouble as a result of his supernormal ability to stop quickly.

Profile graphs (Figs. 4 and 5, pp. 67, 68) speak for themselves in showing the combination of defects that contribute to accidents. A careful inspection will also show the opportunities for clinical supervisors to diagnose and treat ailing drivers.

BIBLIOGRAPHY

Articles by Harry R. DeSilva

Bibliography on Driving Skill, Monograph, *Harvard Traffic Bureau*, 157 pp., 1937.

The Influence of Moderate Carbon Monoxide Poisoning upon the Ability to Drive Automobiles, *The Journal of Industrial Hygiene and Toxicology*, Vol. 19, No. 10, pp. 598–603, December, 1937.

Clinical Treatment of Traffic Violators, *Police Journal*, Vol. XXIV, No. 2, pp. 3–7, December, 1937.

Control of the Driver, *Inspection News, Retail Credit Company*, Vol. 27, No. 1, pp. 11–12, January, 1938.

Facts about Automobile Drivers, *Harvard Alumni Bulletin*, Vol. XL, No. 14, pp. 448–452, January 21, 1938.

Human Element in Stopping a Car, *Brake Service*, Vol. VIII, pp. 183–186, February, 1938.

Driver Testing Results, Monograph, Harvard Traffic Bureau, 100 pp., 1937.

THE RELATIONSHIP OF MEASURES OF VISUAL ACUITY AND AMETROPIA TO READING SPEED*

ELEROY L. STROMBERG
Oklahoma Agricultural and Mechanical College

RECENT experiments dealing with the problem of reading ability have placed a strong emphasis on the physiological causes of reading disability. A number of experimenters have presented data which correlate reading ability and visual efficiency. Many of these reported studies of ametropia and visual acuity in relation to reading have their origin in the work of E. A. Betts. As the result of early investigations, Betts, with the aid of optometrists, constructed the "Betts Ready to Read Test." This test, a series of stereoscopic slides designed for use with the Keystone Ophthalmic Telebinocular, permits the study of the binocular coordination required in reading. Wagner (9) has investigated older children with the Betts technique to show a maturational trend. Witty and Kopel (10, 11) and Swanson and Tiffin (7), however, using the Betts apparatus with elementary school children and college students respectively do not find a relationship between reading ability and measures of acuity and ametropia. Using other visual tests Blake and Dearborn (4), Farris (5), and Fendrick (6), indicate, as the result of studies of college students, high-school students and grade-school children, respectively, that visual acuity and ametropia are clearly related to reading ability.

The literature in this field may be divided into two categories with the work of Betts (1, 2, 3), Wagner (9), Blake and Dearborn (4), Farris (5) and Fendrick (6) stating that errors of

* From the Psychological Laboratory, University of Minnesota.

the refractive type are probably the causes of a large proportion of reading deficiency. Witty and Kopel (10, 11) and Swanson and Tiffin (7) subscribe to the opposite view and show that there is no relation between a student's ability to read and his visual acuity or refractive efficiency.

In the present experiment the relationship between reading speed and visual efficiency, as measured by the Betts tests of Visual Sensation and Perception and the Ives test of Acuity and Ametropia, is investigated.

A group of fast and a group of slow readers were selected from among 1200 University of Minnesota students of Sophomore and Junior rank. All of these students scoring above Q_3 (23 paragraphs) on the Chapman-Cook Speed of Reading Test, Form A were considered in the "fast reader" group. Those whose scores were below Q_1 (16 paragraphs) were considered in the "slow reader" group. These two groups although separated by only seven paragraphs may be considered as two different groups since the standard error of measurement, using the lowest reliability coefficient for the Chapman-Cook test reported by Tinker and Paterson (8), is 2.31 paragraphs. The probability, therefore, that a person who falls in the slow reader group on the first test would fall in the fast reader group on a subsequent test is very small.

It is well established that the level of intellectual ability is an important determinant of reading proficiency. It is necessary then, in selecting subjects for an experiment of this nature, to rule out experimentally the factor of intelligence as a cause of slow reading. All freshmen entering the University of Minnesota take the Minnesota College Aptitude Test which is essentially a test of intellectual aptitude. Arbitrarily the fiftieth percentile of this test was selected as the lower limit of ability for subjects in this experiment. A strict matching technique was employed. Pairs of fast and slow readers were matched on the basis of their College Aptitude Test Centile Scores. In no pair was the difference more than seven centile points. In contrast the smallest difference between the centile

scores of any pair on the Chapman-Cook Speed of Reading Test is fifty centiles. The average difference between the College Aptitude Test centiles for the 71 matched pairs (142 subjects) in the final experiment is 1.41, whereas the average difference between the Chapman-Cook centile scores is 68.295. There are then, in this experiment, a group of fast and a group of slow readers matched for verbal intelligence.

Each of the 142 subjects was tested individually. In the testing procedure the following rule was established and followed rigidly—if the subject wore glasses while taking the Chapman-Cook test, all the succeeding tests were also taken with the correction. The subjects were asked at the time of the reading test to indicate on the blank whether or not they were wearing glasses. In this experiment we are not concerned with the number of individuals who are already wearing a visual correction, nor with the type of correction worn. Whether or not the individual has a visual correction, he remains a fast or slow reader on the basis of his performance on the Chapman-Cook Speed of Reading Test.

The Betts tests of Visual Sensation and Perception, which are described in detail by Betts (2, 3), were administered to each subject. These tests include measures of gross binocular vision, binocular acuity, monocular acuity for the right and left eye, stereopsis (perception of the third dimension), near and far point fusion and ametropia (errors of refraction). Additional measures of binocular acuity (that is, in the normal situation of having both eyes open) were made with the Ives visual acuity test. An additional check of ametropia was made using the cobalt glass test in the Ives apparatus. This apparatus is described in the Bausch and Lomb catalogue.

RESULTS AND DISCUSSION

The Betts Test of Visual Sensation and Perception

This test is administered by presenting stereoscopic slides in the Keystone Ophthalmic Telebinocular. Slides numbered DB 1, DB 2, DB 3, DB 4, DB 5, DB 6, DB 7a and 7b, DB 8, and

DB 10, as indicated in Table I, were used in this experiment. The number of fast and slow readers making each response is given in Table I for each of the sixteen test situations. In seven of these situations the frequency is in favor of the slow readers; in seven the frequency is in favor of the fast readers, and in the remaining two instances the frequencies are equal for the two groups.

TABLE I

Data from the Betts Ready to Read Test Used with the Keystone Telebinocular

(N = 71 fast readers; 71 slow readers)

	FAST READERS	SLOW READERS
ACUITY: Number with less than 100% A.M.A. and Snellen acuity		
Binocular Slide DB 1	22	21
Left Eye Slide DB 2	40	41
Right Eye Slide DB 3	33	29
STEREOPSIS: (third dimension) Number with less than 100%		
Correct responses. Slide DB 6	36	29
FUSION: Number with unquestionable fusion ability		
Far point Slide DB 4	49	48
Near point Slide DB 5	40	45
Hyperopia Slides DB 7a, 7b	26	32
Myopia Slides DB 7a, 7b	23	18
Astigmatism Slides DB 7a, 7b	47	49
Hyperopia plus astigmatism	14	16
Myopia plus astigmatism	21	14
Emmetropia	24	24
Left-eye myopia	16	14
Right-eye myopia	15	15
Left-eye astigmatism	35	36
Right-eye astigmatism	37	43

In order to determine whether such differences as are present are statistically significant, the critical ratio of the frequency proportions[1] for the two largest differences has been calculated. In the test for stereopsis (third dimension perception) the number of cases with less than 100 per cent perception is 29 or 40 per cent for the slow readers and 36 or 51 per cent for the fast readers. The $\frac{D}{\sigma \text{ diff.}}\left(\frac{.11}{.082}\right)$ yields a critical ratio

[1] Yule, G. V., *An Introduction to the Theory of Statistics*, 1924, p. 269, Formula 6.

of 1.34. Similarly, in the case of the combination of myopia and astigmatism, the critical ratio of the difference between the two reading groups is 1.37. These ratios are both in favor of the slow readers but their significance is negligible since they are but slightly above the differences expected from chance.

The results above indicate that the Betts Visual Tests of Physiological Readiness do not differentiate fast from slow readers at the college level. One can scarcely say that poor vision on these tests is related to slow reading since the differences favor the slow readers as often as they do the fast readers. The two largest differences, even though they are not statistically significant, show that the slow readers have a more efficient visual mechanism than the fast readers. This finding is substantiated by the work of Witty and Kopel (10) on elementary-school children. Blake and Dearborn (4) and Farris (5) also find a larger proportion of myopic astigmatism among good readers than among poor readers. It hardly seems logical to attribute poor reading ability to the *absence* of certain visual defects.

The fusion tests show no significant differences between fast and slow readers. These tests are subject to large errors of measurement. The frequencies reported in Table I indicate the number of subjects whose responses are clear-cut and almost immediate. It was noted that if the fusion slides were left before the eyes of the subject for an indefinite period of time, some subjects whose first report indicated a lack of fusion would report a fused image. The time element was not controlled and it seems altogether possible that had all of the subjects been given unlimited time they would all have reported fused images. The data of Witty and Kopel (10, 11) show the same possibility of error in measurement when the test is used with children in the elementary grades.

Table I lists the number of fast and slow readers whose measured responses on the tests of ametropia (refractive anomalies such as hyperopia, myopia, and astigmatism) are different from the normal response. These are given for both the monocular and binocular conditions of vision. The differ-

ences between the two reading groups on all of these tests are far from being statistically significant. It is possible, however, with the Keystone apparatus, to determine the number of diopters of correction necessary to bring the individual's vision up to normal. These measurements are reported in Table II.

TABLE II

Diopter Correction for Ametropia on the Keystone Calibrated Scale

	FAST READERS (N = 71)	SLOW READERS (N = 71)
Myopia		
Left eye	.2008	.1303
Right eye	.2252	.1589
Astigmia		
Left eye	.3812	.3087
Right eye	.3929	.3985

If the fast and slow readers are compared upon the basis of the average amount of correction required, there are again no significant differences. The largest difference is .0725 diopters. Such a difference has no significance since the shaft of the Telebinocular is graduated in .25 of a diopter and the difference is less than the error of measurement. Three of the reported differences (Table II) are in favor of the slow readers.

The Ives Acuity and Ametropia Tests

The Ives screen grid test was used to measure binocular visual acuity. All of the readings were made at twenty feet. The Bausch and Lomb instrument which includes the Ives screen grid test of acuity is calibrated in steps of 0.10 from zero to 2.00. A response of 1.00 is considered normal. For the 139 subjects measured on this test, it is surprising to note that even though the range of readings is from zero to 1.90, only 22 subjects have a reading (the average of four repetitions) of 1.00 or above. The mean acuity reading for the 69 fast readers is .79 and for the 70 slow readers is .73. A difference of .06 between the two means is not significant when one

considers that the calibrations are in units of .10 and the errors of measurement are probably as large or larger than .06.

Since there is apparently no difference between the mean scores of fast and slow readers on the Ives Acuity test, an attempt was made to determine by means of the Chi Square (χ^2)[2] technique whether or not there is any association between the two reading groups and the visual acuity reading.

TABLE III

Derivation of Chi Square (χ^2) for Fast and Slow Readers on the Ives Acuity Test

Ives Acuity	Fast Readers a_1	Slow Readers a_2	$a_1 + a_2$	p.	$a_1 p$
100–190	14	8	22	.636	8.90
90–99	10	10	20	.500	5.00
80–89	15	17	32	.468	7.02
70–79	12	10	22	.545	6.45
60–69	5	13	18	.277	1.38
50–59	3	1	4	.750	2.25
40–49	4	3	7	.571	2.28
0–39	6	8	14	.428	2.56
Totals	69	70	139		35.84

$\chi^2 = 6.6096$ $\quad \bar{p} = .496$

$p > .30$ $\quad \bar{q} = .494$

Table III presents the data from which the χ^2 was determined. The probability obtained from the observed χ^2 value is greater than .30. This indicates that the association can be accounted for in terms of errors of random sampling. From this we may conclude that fast and slow reading, as represented in our sample, are not associated with visual acuity as measured by the Ives test. This finding disagrees with that of Fendrick (6) who concludes that his results add much weight in establishing the significance of visual acuity as a factor affecting reading efficiency. His study was made on children and that may account for the differing results.

In the present experiment ametropia was measured by two methods: (1) The Betts Test, and (2) The Ives Test. The

[2] Fisher, R. A., *Statistical Methods for Research Workers*, Oliver and Boyd. London, 1936.

Betts test involves visual discriminations at the twenty foot and infinity equivalents on the Keystone Telebinocular. The Ives test involves a differential response from the distance of twenty feet to a light shining through a cobalt blue glass. For an individual with normal vision the glass appears to be lavender, for the myopic individual it is red with a blue border, for the hyperopic it is blue with a red border. In order to determine whether the two tests of ametropia were measuring similar functions, a contingency coefficient and χ^2 were calculated for the distributions. The probability obtained from the observed χ^2 value is less than .01. This indicates that the two measures are associated and that the association cannot be accounted for in terms of errors of random sampling. This association in terms of the contingency coefficient is .449. The Betts and Ives test do measure somewhat the same functions. Which is the criterion of validity cannot be stated. If the Betts test is a valid test of ametropia, then the validity of the Ives test is not high. The reverse can be stated just as truly.

CONCLUSIONS

The results of this study indicate that the Betts Tests of Visual Sensation and Perception do not differentiate fast from slow readers at the college level. Differences in visual efficiency are in favor of the slow reader as often as they are in favor of the fast readers. Optometrical measurements of acuity, ametropia, fusion and stereopsis are not related to reading efficiency as measured by the Chapman-Cook Speed of Reading Test.

There is no difference between the two reading groups on the Ives Acuity test. Statistical treatment of the results shows that any association between the acuity scores and reading speed can be accounted for in terms of errors of random sampling.

The data of this experiment are of necessity confined to those functions which one can observe by the use of various measuring instruments. These functions are spoken of as "peripheral factors" and deal primarily with the efficiency of the sense organ in question. Our data lend no support to the "periph-

eral factor'' theory of reading disability which has received so much emphasis recently. In no one of our measurements of visual efficiency is there a definite and unquestionable relationship to reading ability. Individual case studies and the treatment of the group data indicate that so far as reading speed is concerned acuity and emmetropia are of little consequence. The data presented in this report, therefore, definitely contradict the contention that certain peripheral factors, namely acuity and ametropic errors, are important determinants of reading proficiency.

It seems probable, therefore, that we must look to the so-called ''central factors'' such as perceptual habits, proficiency of comprehension and assimilation, apprehension span, and the like for the important determinants of reading ability.

BIBLIOGRAPHY

1. BETTS, E. A. ''A Physiological Approach to the Analysis of Reading Disabilities,'' *Ed. Res. Bull.*, 1934, 13, 135–140, 163–173.
2. ————. *Manual of Directions for Ready to Read Tests.* Meadville, Pa.: Keystone View Co., 1934, pp. 48.
3. ————. *The Prevention and Correction of Reading Difficulties.* Chicago, Ill.: Row, Peterson & Co., 1936.
4. BLAKE, M., and DEARBORN, W. F. ''The Improvement of Reading Habits.'' *J. Higher Ed.*, 1935, 6, 83–88.
5. FARRIS, L. P. ''Visual Defects as Factors Influencing Achievement in Reading.'' *J. Exper. Educ.*, 1936, 5, 58–60.
6. FENDRICK, P. *Visual Characteristics of Poor Readers.* Teachers Coll. Contr. to Educ., No. 656, 1935, pp. 54.
7. SWANSON, D. E., and TIFFIN, J. ''Betts Physiological Approach to the Analysis of Reading Disabilities as Applied to the College Level.'' *J. Educ. Res.*, 1936, 29, 433–448.
8. TINKER, M. A., and PATERSON, D. G. ''Studies of Typographical Factors Influencing Speed of Reading. XIII. Methodological Considerations,'' *J. Appl. Psychol.*, 1936, 20, 132–145.
9. WAGNER, G. ''The Maturation of Certain Visual Functions and the Relationship Between These Functions and Success in Reading and Arithmetic.'' *The Phi Delta Kappan*, 136, 27, 663–671.
10. WITTY, P. A., and KOPEL, D. ''Factors Associated with the Etiology of Reading Disability.'' *J. Educ. Res.*, 1936, 29, 449–459.
11. ————. ''Heterophoria and Reading Disability.'' *J. Educ. Psychol.*, 1936, 27, 222–230.

THE ESTIMATION OF PERFORMANCE-TIME IN SIMPLE TASKS

GEORGE J. DUDYCHA AND MARTHA M. DUDYCHA
Ripon College

I. INTRODUCTION

SINCE the beginning of the present century investigators have been periodically interested in the problem of time estimation. MacDougall (6) in 1904 directed his attention to the sex differences in the "sense of time." Yerkes and Urban (13) inquired into the relation of time-estimation to age, sex and physiological rhythm. Myers (7) published his study of "incidental perception," later followed by Spencer's (8) article on the effect of different interpolations on time estimation. Axel (1) determined the effect of different fillings on the estimation of the duration of time, and the influence of age on this ability. Swift and McGeoch (9) compared the perception of filled and unfilled time, whereas Gulliksen (3) studied the influence which the particular task has on the perception of time. Among the more recent students of this problem are: Woodrow (12), Whitely and Anderson (11), Triplett (5), and Hawickhorst (4).

In the above studies the central purpose was to ascertain with what accuracy subjects *estimated elapsed time* when certain factors were systematically varied and controlled. In the present study[1] the procedure was reversed; the subject was assigned a simple task and asked to *estimate his performance-time before beginning the task*. This is perhaps the more practical situa-

[1] The present study was originally performed as a part of an extensive study of punctuality. The relationship between the prediction of performance-time and punctuality is presented in the following monograph: Dudycha, G. J.: "An Objective Study of Punctuality in Relation to Personality and Achievement," *Arch. of Psychol.*, 1936, No. 204.

tion, for we are probably more often called upon to estimate performance-time than to judge elapsed time. Thus the general purpose of this study is to answer three questions, namely: How accurately is the performance-time for certain simple tasks predicted? How do the sexes differ in such prediction? Is there any evidence of the presence of a general ability to predict performance-time?

II. PROCEDURE

The subjects in this experiment were 185 college students, of whom 109 were men and 76 women. They were asked to perform six simple tasks, which were selected because of the ease with which they could be performed, their probable duration, and their similarity to common tasks frequently performed. They were: (1) *Assembling letters:* Letters on small, cardboard squares were arranged to spell the words given. Five different lists of words were used (the same 65 letters were included in each set), but each student used only one list. (2) *Obtaining sums:* Each student was given one of four sets of addition problems to solve. (3) *Copying geometric figures:* Geometric figures consisting of straight lines arranged to form complex figures drawn on cross-section paper were copied (not traced) on another sheet of cross-section paper. (4) *Assembling a jig-saw puzzle:* A simple jig-saw puzzle consisting of 28 pieces was assembled with the help of a picture. (5) *Running a non-social errand:* When sent on a non-social errand, the student was required to retrieve a card bearing a certain number which was previously planted in an accessible but more or less distant place, as on a certain shelf in the library, or in the drawer of a table in a certain room, et cetera. Before the student went on his mission, the writers made certain that the student knew where the card was planted. (6) *Running a social errand:* In the social errand, the students were sent on missions which definitely involved another person; such as, withdrawing a book from the library, asking for a circular in a shop, buying small articles in stores—particularly groceries. In each case, the student made his estimate before the task was begun, and was

requested not to time himself. The experimenters recorded all estimated and performance times in half-minutes and kept them secret until the entire experiment was completed.

Since both the estimated and performance times varied considerably from person to person and from task to task, it was evident that a score, as a ratio, was necessary as a common basis for making comparisons. Thus the prediction-of-performance-time score was obtained as follows: First, the predicted-time was divided by the performance-time. All overestimation was designated as plus and underestimation as minus. And finally 1 was subtracted from each overestimation score, and all underestimation scores were subtracted from 1. Thus equal over- and underestimation give comparable scores—numerically equal but differing in sign.[2]

III. RESULTS

According to the medians of Table I we see that the performance-time was, on the average, overestimated in five of the six tasks, and that the one task which was underestimated was the reproduction of geometric figures. Further, we observe that the time for assembling the jig-saw puzzle was overestimated most and that for running the social errand least. Perhaps the reason for the small overestimation of the latter is the students' greater familiarity with this type of task. Finally, the medians indicate that, on the average, the error of estimation was not great in any of the six situations, being from about one-fourth to less than one-half greater than the actual performance-time. These errors are not large in terms of minutes when we remember that the performance-times were also short. The general tendency toward overestimation may be accounted for by the attitude frequently expressed by the subjects in words similar to the following: "I had better allow myself enough time, since I do not know just how long it will take me." Many of the students seemed to feel that they were allowing more than enough time when estimating the geometric

[2] For a more complete discussion of score determination see *Arch. of Psychol.*, 1936, No. 204, pp. 39–40.

TABLE I

Medians and Interquartile Ranges for the Prediction of Performance-Time

TASK	MEDIAN*	Q
Assembling letters	.340	.500
Obtaining sums	.356	.436
Copying geometric figures	−.233	.245
Assembling jig-saw puzzle	.412	.445
Running non-social errand	.372	.311
Running social errand	.259	.281

* In terms of scores scaled from zero.

figures, even though most of them underestimated it. Perhaps lack of familiarity with this particular task, or lack of appreciation of its details, accounts for their rather consistent underestimation.

Accuracy in estimating the time required for these tasks would probably have increased had several trials been taken, but this was not the purpose of the study. The authors sought to find with what accuracy students estimate the duration of simple tasks, similar to those performed in daily life, without preliminary or specific practice. This is a common life situation. Quite constantly we are called upon to estimate the length of time necessary to accomplish tasks which, although partly similar to others we have performed, are novel. To duplicate this situation, only one trial was given in each task. Both Koehnlein (5), and Hawickhorst (4) report improvement in estimation with practice, but also observe that such improvement is usually temporary.

Marked sex differences in the estimation of time have been reported by MacDougall (6), Yerkes and Urban (13), Myers (7), Axel (1), and Gulliksen (3). Each found that women overestimate far more than men do. This observation is not borne out by our data. As may be seen in Table II, the six medians for the women do indicate that they allow themselves, on the average, slightly more time than the men do, but the differences between the sexes are not reliable, as the critical

ratios indicate. The greatest difference is in assembling the jig-saw puzzle, in which case there are 89 chances in 100 of a true difference; the smallest difference, 51 chances in 100, is in running the social errand. Since none of the critical ratios is reliable, we cannot conclude that most women usually allow more time for performing tasks than men allow.

TABLE II

Reliability of the Difference Between Men and Women in Ability to Predict Performance-Time

	SEX	MEDIAN	Q	CRITICAL RATIOS	CHANCES IN 100
Assembling letters	Male Female	.310 .360	.512 .465	.55	64
Obtaining sums	Male Female	.321 .450	.433 .444	1.57	85
Copying geometric figures	Male Female	−.250 −.209	.249 .285	.80	71
Assembling jig-saw puzzles	Male Female	.330 .471	.455 .388	1.83	89
Non-social errand	Male Female	.346 .400	.311 .275	.98	75
Social errand	Male Female	.258 .260	.268 .297	.04	51

A second sex difference that has been observed by a number of investigators is that women vary more in their judgments than do men. This, however, is not found in our results, as is indicated by the Q's presented in Table II. In three of the tasks the men have larger Q's, and in the other three tasks the women have the larger Q's. The men varied more in their estimates of assembling letters and the jig-saw puzzle, and in running the non-social errand, whereas the women varied more in their estimates in obtaining sums, drawing geometric figures and running social errands.

Finally, there is no greater preference on the part of women for numbers ending in 0 or 5, or even rather than odd numbers, as has been reported by other students. Very slight differences

between the sexes were found in this regard; both showed about equal preferences for the various numbers.

The third question raised in this paper is with regard to the presence of a general ability to predict performance-time. To test this hypothesis, the Chi-square method, developed by R. A. Fisher (2), was employed. Chi square indicates the presence of an association between two sets of data, but it does not indicate the degree of such an association. The Chi-square method was used, rather than the Pearson *r*, because it is based on broad categories rather than on fine measures used to obtain a coefficient of correlation. A 5×5-fold classification was used. The limits for the five classes are as follows: greatly overestimated, above $+.5$; overestimated, between $+.5$ and $+1$; correct estimation, between $+.1$ and -1; underestimated, between $-.1$ and $-.5$; greatly underestimated, above $-.5$. Since a 5×5-fold

TABLE III

Measures of Association between Estimates in Different Types of Situations

	ASSEMBLING LETTERS	OBTAINING SUMS	COPYING GEOMETRIC FIGURES	ASSEMBLING JIG-SAW PUZZLE	NON-SOCIAL ERRAND
Obtaining sums					
Chi square	9.7				
P-value	.87				
Copying geometric figures					
Chi square	33.2	23.1			
P-value	.01	.11			
Assembling jig-saw puzzle					
Chi square	13.3	.6	21.4		
P-value	.64	.99	.17		
Non-social errand					
Chi square	10.1	16.9	.4	10.3	
P-value	.85	.40	.99	.85	
Social errand					
Chi square	4.3	13.1	14.5	16.6	17.4
P-value	.99	.64	.57	.44	.37

classification had been used, Fisher's table had to be entered with n equal to 16 in order to find the P-values which correspond to the Chi squares. The 15 Chi squares with their corresponding P-values are given in Table III. Because Fisher has arbitrarily decided that there is no significant association unless P is .05 or less, we have found only one significant association; namely, that between assembling letters and drawing geometric figures. Since 14 of the 15 Chi squares do not indicate a significant departure from independence, our conclusion from the present data must be that, as far as these particular types of situations are concerned, there is no general ability to predict performance-time. In other words, if a student overestimates his performance time in one of these tasks, we have no reason to suppose that he will be equally inaccurate in estimating the time in another.

IV. CONCLUSIONS

1. The students overestimated their performance times in five of the six tasks used, but the error in each of the situations was small, being on the average, between 25 and 41 per cent.

2. Although the women on the average allowed themselves more time for completing the tasks than the men did, there is no reliable difference between the sexes in this respect.

3. Women were not found to be more variable in their judgments than men, nor did women show any greater preference than the men for certain numbers.

4. The results do not reveal the presence of a general ability to predict performance-time.

REFERENCES

1. AXEL, R. "Estimation of Time," *Arch. of Psychol.*, 1924, No. 74.
2. FISHER, R. A. *Statistical Methods for Research Workers* (5th ed.) 1934.
3. GULLIKSEN, H. "The Influence of Occupation upon the Perception of Time," *J. of Exp. Psychol.*, 1927, 10, 52–59.
4. HAWICKHORST, L. "Mit welcher Sicherheit wird der Zeitwert einer Sekunde erkannt?" *Zsch. f. Sinnesphysiol.*, 1934, 65, 58–86.
5. KOEHNLEIN, H. "Uber das Absolute Zeitgedächtnis," *Zsch. f. Sinnesphysiol.*, 1934, 65, 35–57.

6. MacDougall, R. "Sex Differences in the Sense of Time," *Science*, n.s., 1904, 19, 707–708.
7. Myers, G. C. "Incidental Perception," *J. of Exp. Psychol.*, 1916, 1, 339–350.
8. Spencer, L. T. "An Experiment in Time Estimation Using Different Interpolations," *Am. J. of Psychol.*, 1921, 32, 557–562.
9. Swift, E. J., and McGeoch, J. A. "An Experimental Study of the Perception of Filled and Empty Time," *J. of Exp. Psychol.*, 1925, 8, 240–249.
10. Triplett, D. "The Relation between the Physical Pattern and the Reproduction of Short Temporal Intervals: A Study in the Perception of Filled and Unfilled Time," *Psych. Monog.*, 1931, 41, No. 4, *Univ. Iowa Stud. Psychol.*, No. 14, 201–265.
11. Whitely, P. L. and Anderson, J. C. "The Influence of Two Different Interpolations upon Time Estimation," *J. of Gen. Psychol.*, 1930, 4, 291–401.
12. Woodrow, N. "Behavior with Respect to Short Temporal Stimulus-Forms," *J. of Exp. Psychol.*, 1928, 11, 167–193.
13. Yerkes, R. M. and Urban, W. M. "Time Estimation in Its Relation to Sex, Age and Physiological Rhythms," *Harvard Psychol. Studies*, 1906, 2.

COMPARISON OF THE STANDARD AND A "SELF-SCORING" FORM OF THE VINELAND ADJUSTMENT SCORE CARD

S. JEAN WOLF
New York University

I

A COMPARATIVE study[1] was made of two groups of girls of relatively equal intelligence but differing markedly in academic achievement. The two groups, 50 in each, were equated as to: (1) normal intelligence, (2) normal health, (3) Italian nationality, (4) American nativity, (5) residency in a typical Italian section of New York City, (6) school grade, 6A or 6B, (7) age, 11 to 12 years.

The two groups differed markedly in academic achievement, namely, one group succeeded and the other failed in the standard school subjects as measured by the Metropolitan Achievement Tests.

The groups were compared primarily as to whether they manifest measurable differences in such phases of personality as a representative set of selected personality tests purports to measure.

The Vineland Adjustment Score Card was one of a group of personality measures used in this study. This score card is a social adjustment scale, devised by L. N. Yepsen,[2] based upon clinical analysis of individual deviates in a large group of children. It is reported to be a valid and reliable measure

[1] Wolf, S. J.: A Comparative Study of Two Groups of Girls of Relatively Equal Intelligence but Differing Markedly in Achievement. Ph.D. Thesis, New York University, 1936. New York University, Library. See abstract, JOURNAL OF APPLIED PSYCHOLOGY, June, 1937.

[2] Yepsen, L. N.: Measuring Social Adaptation. Journal of Juvenile Research, 12, Sept.–Dec., 1928, pp. 254–260.

of social adaptation consisting of 14 major categories dealing with certain phases of adjustment of the individual to his environment. Under each of these categories are 4 to 7 statements of observable behavior, there being a total of 70 such statements. Each describes a possible characteristic action of the individual being scored. For example, under the first general category, "Attitude of Others Toward Him," the following are the first and the last of the seven characteristics:

"Choose him as a leader."

"Butt of crowd, pick on him."

The person scoring checks one statement under each category that is most true of the individual rated. Each statement is weighted for scoring. The scores of the hypothetical average individual is 100, this being the modal activity of the 14 items statistically weighted.

In this investigation, each girl's social behavior was scored on the basis of the Vineland Adjustment Score Card by her teacher in whose class she had been at least one full term. The two experimental groups were compared on the basis of the total scores.

In Table I are indicated the statistical values of this comparison.

TABLE I

Statistical Values of Vineland Adjustment Score Card, Total Scores, Scored by Teachers

	SUCCESS GROUP 50 CASES	FAILING GROUP 50 CASES
Range	26–137	43–125
Mean	97.38 ± 1.45	91.18 ± 2.04
Median	102.5 ± 1.81	97.0 ± 2.55
S.D.	15.23 ± 1.02	21.35 ± 1.44
Difference of Means	6.20 ± 2.5	
Critical Ratio	2.48	

The Vineland Score Card was further used in the following way: Each statement in the scale definitizing a type of behavior was recast into the first person, retaining the standard form,

categoric grouping and sequence. For example, "Choose him as leader," was changed to "I am chosen as leader," "Butt of crowd, pick on him," was changed to, "I am the butt of a crowd, they pick on me."

Each girl scored herself on the "self-scoring" form of the Score Card.[3] In Table II appear the statistical values of the "self-scoring" form.

TABLE II

Statistical Values of "Self-Scoring" Form of the Vineland Score Card, Total Scores, Scored by Students

	SUCCESS GROUP 50 CASES	FAILING GROUP 50 CASES
Range	56–128	37–142
Means	99.64 ± 1.45	92.52 ± 2.14
Median	102.5 ± 1.81	95.5 ± 2.11
S.D.	16.80 ± 1.02	22.44 ± 1.51
Difference of Means	7.12 ± 2.39	
Critical Ratio	2.67	

The value of self-rating scales is often questioned, as the subject rating himself may be inclined to hide such truth as may not be complimentary to himself, thus over-rating himself. Fear of poor grades is also considered to influence test performance. With these factors in mind care was exercised in the administration of all tests in this study to eliminate all possible fears and self-defense attitudes. Through examples, assurances, and suggestions, an effort was made to establish attitudes of honesty and frankness.

A comparison of values in Tables I and II does not appear to indicate that any fears or deceptions affected the self-scoring. It would seem, on the contrary, that the children rated themselves critically and honestly. The differences between the groups scored by the teachers (Table I) and by the children (Table II) are closely alike. In both instances the difference was in favor of better adjustment for the success group.

It is of interest to note the mean scores of the two groups on

[3] The findings of the self-scoring were not included in the original report.

the standard form and on the "self-scoring" form. All the means are markedly close to 100, which is the score of the hypothetical average individual adopted in the standardization of the scale. These findings appear to speak well for the standardization and validity of the Score Card.

It is also of interest to note the ranges, standard deviations and median scores, as they relate to the experimental groups. The differences between the standard deviations on both forms for the respective groups are negligible. Likewise, differences between the medians are either absent or negligible. Thus it appears that the central tendencies for the two groups on the standard form and the "self-scoring" form are closely alike.

The Telling What I Do Test[4] was another personality measure administered in this study. This test purports to measure social, ethical and moral standards of behavior. It obtains statements concerning their attitudes towards a variety of situations which children commonly experience. With respect to the purpose of the measure and the experiences investigated, the Telling What I Do Test is more like the Vineland Score Card than any other personality measure used in this study. The Telling What I Do Test differentiated the two groups, in terms of the Critical Ratio, to the extent of 3.27.

The scores obtained on both forms of the Vineland Score

TABLE III

Coefficients of Correlation of Scores on the Teachers' and Self-rating Forms of the Vineland Score Card and the Telling What I Do Test for the Success and Failing Groups

	SUCCESS GROUP	FAILING GROUP
Teachers' and Self-Rating Forms of Vineland Score Card	.21 ± .09	.17 ± .09
Vineland Score Card Teachers' Rating and Telling What I Do	.43 ± .07	.12 ± .09
Vineland Score Card Self-Rating Form and Telling What I Do	.56 ± .06	.50 ± .07

[4] By Harry J. Baker, Published, Public School Publishing Co., Bloomington, Ill., 1930.

Card were correlated with each other and with the scores of the Telling What I Do Test, for each group. The coefficients of correlation are presented in Table III.

An examination of the figures in Table III indicates: (1) that the extent of agreement between the teachers' and self-rating on the Vineland Score Card is relatively low for both groups ($.21 \pm .09$ and $.17 \pm .09$), (2) that three of the four coefficients of correlation between the teachers' rating and the children's self-ratings are markedly low ($.21 \pm .09$, $.17 \pm .09$, $.12 \pm .09$), (3) that the highest coefficient of correlation between the teachers' ratings and the children's self-ratings, namely, $.43 \pm .07$, is for the success group which is the better adjusted group and in which case the adjustment characteristics are more easily discernible, (4) that the teachers show relatively marked disagreement with the children in every instance where their ratings are correlated, although both the teachers and children agree as to relative maladjustment of the failing group on the basis of central tendencies. They evidently disagree to a large extent as to the adjustment of individual children and as to particular items that point to maladjustment.

II

An item analysis was made of the standard form of the Vineland Adjustment Score Card, scored by the teachers, and of the "self-scoring" form where each girl scored herself. The purpose of making this analysis was to ascertain on what particular characteristics reported in the Score Card, the success and failure groups differed most or least on each of the Score Card forms, and to ascertain the degree of statistical reliability of such possible differences between the two groups.

Both forms of the Score Card were item-analyzed in terms of the percentage of children in each of the groups that checked each of the seventy statements which constitute the scale. The differences and probable errors of the differences of the percentages in each group that checked each particular statement in the scale were obtained and the critical ratios computed.

In Tables IV and V respectively, are presented the fifteen statements in the standard form of the Vineland Score Card checked by the teachers which distinguished most and least, respectively, between the success and failing groups in terms of critical ratios. The starred critical ratios in the tables indicate that the greater number of checks for the particular statement was for the failing group. The columns at the right indicate the number of checks in each group for each particular statement. The seventy definitizing statements in the scale are not numbered. The "self-scoring" form is exactly the same in content as the standard form, but each statement is presented in the first person. The statements in this form likewise are not numbered. The numbers appearing to the left of the statements in the tables are used to facilitate reference and comparison in this article.

TABLE IV

Fifteen Statements of the Vineland Adjustment Score Card Checked by the Teachers that Differentiate Most Between the Success and Failing Groups

	C.R.	SUC. GR.	FAIL. GR.
1. Begins new steps readily upon suggestion	5.83	31	14
2. Can be left to own resources	4.18	20	8
3. Interest in work at hand	3.83	21	10
4. Cooperates, is considered a leader	3.28	4	0
5. Never does more than necessary	3.05*	10	19
6. Works with only occasional direction	2.46	32	24
7. Occasional flare up	2.40*	6	12
8. Always complaining about something	2.32	5	3
9. Makes the most of any situation	2.24*	14	21
10. Adjust self to social situation	2.16	25	18
11. Ignore and shun him	2.15	2	4
12. Play with him only occasionally, not often	2.13*	5	10
13. Most time refuses to work	2.00	1	0
14. Grumbles regarding necessary tasks	2.00*	1	5
15. Enjoys being by self to the extent of being reclusive	2.00	1	0

In Tables VI and VII respectively, are presented the fifteen statements in the "self-scoring" form of the Vineland Adjust-

ment Score Card checked by the children that distinguish most and least between the success and failure groups in terms of critical ratios. The starred figures, the columns to the right and the numbers to the left of the statements have the same significance as in Tables IV and V.

TABLE V

Fifteen Statements of the Vineland Adjustment Score Card Checked by the Teachers that Differentiate Least Between the Success and Failing Groups

	C.R.	SUC. GR.	FAIL. GR.
1. Word not to be relied on	0	3	3
2. Is cruel, a "bully"	0	3	2
3. Profane, vulgar	0	0	0
4. Actually purloins	0	0	0
5. Destroys own and others	0	0	0
6. Claims others property but gives it up	0	2	2
7. Works without direction	.31	12	11
8. Responds to positive measures	.38	8	7
9. Seek his companionship	.54	5	4
10. Ignore and shun him	.60	4	3
11. Falsifies but readily admits falsehood	.60	4	3
12. Always courteous, language good	.60	4	3
13. Must be given only a few explanations	.60	3	2
14. Seldom requires discipline	.66	36	34
15. Sees something to be done and does it in approved manner	.80	6	8

Comparison of Tables IV and VI indicates relative correspondence between the opinions of the teachers and the children as to the personality characteristics (comprised in the Vineland Score Card) that differentiate most between the success and failure groups. Only two statements correspond entirely, namely Nos. 4 and 6 in Table IV and Nos. 2 and 8 in Table VI respectively. Some have close (though not complete) correspondence, namely Nos. 1, 4, 6 in Table IV and Nos. 11, 9, 3 in Table VI respectively. Some statements similar in thought content appear in both tables, but are considered by the teachers as more true of the success group and by the children as more true of the failing group. These are Nos. 3, 4, 11, 15 in Table IV and Nos. 4, 10, 6, 5 in Table VI respectively.

TABLE VI

Fifteen Statements of the Vineland Adjustment Score Card, Self-Scored By the Children that Differentiate Most Between the Success and Failing Groups

	C.R.	SUC. GR.	FAIL. GR.
1. I am rough but mean well	5.83*	4	14
2. I cooperate, am considerate, a leader	5.36	14	2
3. I work with only direction	4.43*	18	32
4. I am interested in work at hand	4.06*	13	26
5. I enter into group play unwillingly	3.88*	2	10
6. I am the butt of the crowd, they pick on me	3.88*	1	8
7. I must be watched continually	3.54*	0	5
8. I work with only occasional direction	3.41	12	4
9. I am accepted readily as one of the group	3.13	18	9
10. I am accepted as a leader	3.10*	2	8
11. I willingly carry out instructions	2.81	25	16
12. I can be left to my own resources	2.74	30	21
13. I am careful of my own but not of others' property	2.66	1	5
14. I must be given frequent attention	2.50*	5	11
15. I am not known to tell falsehood	2.44	27	19

TABLE VII

Fifteen Statements of the Vineland Adjustment Score Card Self-Scored by the Children that Differentiate Least Between the Success and Failing Groups

	C.R.	SUC. GR.	FAIL. GR.
1. I am chosen as a leader	0	7	7
2. I am played with (associated) occasionally	0	6	6
3. I am ignored and shunned	0	2	2
4. I actually purloin others' property	0	1	1
5. I refuse to conform and have to be spoken to frequently	0	3	3
6. I require constant supervision	0	1	1
7. I am careful of my own and others	0	38	38
8. I have an occasional flare up	.28	5	3
9. I make the best of any situation	.30	23	22
10. I try to guide others in recognition of personal rights	.32	16	17
11. I generally play games of individual nature	.35	11	10
12. I work without direction	.40	8	7
13. I require only occasional check up	.42	8	7
14. An initial word is enough	.45	7	6
15. I often become dissatisfied	.50	4	3

Comparison of Tables V and VII indicates that statement 4 in each of these tables is the only adjustment characteristic upon which the teacher and the children agreed as to it being one of the 15 least differentiating between the two groups.

CONCLUSIONS

In the light of the findings in this study, the following conclusions seem logical:

1. There is marked to strong probability that the Vineland Score Card is a valid instrument for differentiating groups of achievement-succeeding from achievement-failing children who are otherwise fundamentally alike (See Tables I and II).

2. Failure in academic achievement is accompanied by inferior adjustment.

3. Special effort to dispell fear of consequences seems to be a factor, particularly with children, in inducing honesty and frankness in self-scoring, and, therefore, in evaluating self-rating tests and scales.

The fact that the self-rating form of the Vineland Score Card differentiated the two groups with somewhat greater reliability (statistically) than the standard form, in the terms of central tendencies, appears to be significant and warrants further experimentation with the self-rating form of this scale.

5. The fact that a representative number of statements (15) on each form of the Vineland Score Card that differentiate the two groups either most or least reliably (statistically) differs so markedly, appears to be significant, and warrants further experimentation with these particular statements and statements similar in nature (See Tables IV, V, VI, VII).

6. The fact that the coefficients of correlation between the teachers' ratings on the standard form of the Vineland Score Card and the children's ratings on the self-rating form of this Score Card and on the Telling What I Do Test, are relatively low (Table III), appears to be significant, and warrants further investigation as to the validity of the teacher-rating and self-rating on these scales.

7. There may be a possibility that the children's estimate of themselves is more valid than the teacher's. The children appear to have been unrestrained in rating themselves and were substantially consistent. Hence the qualities they admitted may possibly be more true than those the teachers attributed to them.

8. More experimentation in the manner here indicated may demonstrate that an honest self-analysis by children may be of greater value than analysis by others. More definite conclusions depend upon further verification of the above findings in other similar studies and especially in widely different environments.

RELATION OF SCORES IN STRONG'S INTEREST ANALYSIS BLANKS TO SUCCESS IN SELLING CASUALTY INSURANCE

MARION A. BILLS
Ætna Life Insurance Company

SINCE early in 1932, the Ætna Casualty and Surety Company has conducted each year, seven five-weeks schools for commissioned Casualty salesmen. These schools are entirely voluntary and the men come at their own expense; however, they must be recommended by an Ætna Manager in the Field and approved by the Home Office Agency Department. The average attendance at these schools is 30 men whose ages and experience have varied from young men just out of college with no selling experience, to mature men with up to 15 years selling experience. They have come from practically all states in the Union and from rural and urban territories.

Since January, 1933, in cooperation with the heads of the school, we have given to each school the Vocational Interest Analysis—E. K. Strong, Jr.; the Personality Inventory—R. C. Bernreuter, and a self-rating blank based on one used by A. W. Kornhauser. These were given, as stated, to each school on an experimental basis and we hoped that from some one of these tests or from some combination of all of them, to be able to predict the selling success of the men well enough so that the test might be used as an aid in the future selection of agents. At the time we started the study we definitely planned it as a five-year program and although the results of the test have been used for the last two years by our Casualty Agency Division, a complete study of results was started in January of this year and we wish here to report on the five years' experience with

the Strong Interest Analysis. Further reports will be made at a later date.

CRITERION—RATINGS ON SUCCESS BY MANAGERS

For a study of this kind we needed, of course, not only the test score but a definite idea of the future progress of the individual. There was some discussion of using as this criterion the actual sales made, but for many reasons this was not feasible and we have therefore taken as our criterion of success the opinion of the managers or general agents under whom the returned school men worked. The five divisions in which the managers were asked to place the men were the outcome of the natural ways of expression when we talked with the heads of the school and the managers:

Outstanding successes	OSS
Successes	S
Fair, will probably remain in business	F+
Poor, will probably leave the business very shortly	F–
Failure, already out of the business	O

We have not asked for this rating of the men until after they had been in the field for a full year following attendance at the school. The present study was based on the managers' ratings made in the summer of 1937 of the men who were in the school during the years 1933, 1934, and 1935, and a few members of the early school of 1936 where the managers felt it possible to give us a definite rating. In most cases this rating was compared with previous ratings and checked by the heads of the school and the Home Office Agency Department. In doubtful cases we corresponded further with the manager. In almost all instances the managers gave us a short memorandum concerning each man as well as their estimate into which of the five classes he fell.

In certain instances the managers gave us a memorandum only and we have used our best judgment in classifying these. For instance, a person concerning whom the memorandum said, "About a year ago married a wealthy woman. Has not

been near the office since and has produced nothing. Might be able to sell if he had to" was classed as a failure, while a person concerning whom a memorandum said, "Took salary job on account of expense of wife's illness. Expect he will return to the business" was classed as undetermined. To simplify this study we have combined those marked successes and fair plus and on the low side those marked probable failures and failures.

There was every indication that the managers took the work very seriously and gave us their best considered rating of the individual.

NUMBER OF CASES

The total number of cases considered in the group was 744. However, 27 of these had been sent from banks or offices with no intention of going into sales upon their return. These we classed as inside men and felt that they should not be included in the study. Besides these we had 30 who were placed in an undetermined group. Most of these were men who had returned to brokerage firms and where, while the manager knew the business from the firms, he had no way of knowing how much business was brought in by the individual persons. There were a few cases placed in the undetermined group as the one quoted above. This leaves us a total of 687. Of these, 99 were definitely returning to family agencies and these we studied in a separate group; that is, for our regular group we had 588 cases, on all of whom we had test scores and a rating of the manager and on a large number of whom the rating of the manager had been checked by the Home Office Agency Department.

INTEREST ANALYSIS

We found the most valuable of the three tests in predicting success was the interest analysis. It has been standardized by Dr. Strong for three definite sales occupations—life insurance sales, real estate sales, and specialty sales. At the beginning we expected to use all three of these occupational standards

but in an early analysis and borne out by the present work, the specialty sales added nothing to the prediction and was therefore dropped, leaving two definite sales occupations to be considered. Other occupational gradings had been tried out but had not added to the predictability of the test and had also been dropped. We therefore considered for this statistical study the scoring on the two sales groups—life insurance sales and real estate sales. The scoring advocated by Dr. Strong is by letter, "A" meaning that the interests of the individual center very closely around the same interests as people successful in that field; "B" meaning that they center somewhat around the same interests; and "C" meaning that they do not center at all around the interests of the people who have been successful in that field. To combine the two scores we have arbitrarily given the following weights to the letters:

A	+3
B+	+2
B	0
B-	-2
C	-3

The question will naturally arise why we did not use the actual numerical scores back of the letter scores. It would have been perfectly possible to have done this, but the primary object of the study was as a guide to laymen who more easily talk in letter ratings than in numbers.

RESULTS

The following table shows for the entire group of 588 the relationship between test score and success ratings by the managers:

From the above table it is evident that without considering any further data than test score we get a difference in the percentage of failures running from 22 to 76, or there is about 3½ times the chance of failure for the low score group than for the high. It is also of special interest to note the consistency of the increase in the percentage of failures with a decrease in score.

SCORE IN INTEREST ANALYSIS *versus* MANAGERS' RATINGS

Casualty School Men—Percentage in Each Group

Key: OSS—Outstanding Success S—Successes and Fair Plus
F—Failures and Probable Failures

All Ages

RATINGS BY MANAGERS	SCORES IN INTEREST ANALYSIS					
	+6	+4 +5	+3 −2	−3 −5	−6	
	%	%	%	%	%	
OSS	25	16	11	8	4	
S	53	56	47	39	20	
F	22	28	42	53	76	
Total No.	139	130	193	71	55	588

This table covers all ages. However, it seemed advisable to break it down into age groups since records show that the chance of success for persons under 25 is considerably less than for those over. In breaking it down into the smaller groups we have, on account of the small number of cases, condensed the two lower score groups into one, so that we might have a sufficient number in each group for percentages to have meaning. Also, so few of those 24 years of age and under were rated by the managers as outstanding successes that we have combined them with the success group. The following tables illustrate the result for the various age groups:

Age 24 and Under

RATINGS BY MANAGERS	SCORES IN INTEREST ANALYSIS				
	+6	+4 +5	+3 −2	−3 −6	
	%	%	%	%	
S	59	67	45	25	
F	41	33	55	75	
Total No.	39	43	71	47	200

Ages 25–29

RATINGS BY MANAGERS	SCORES IN INTEREST ANALYSIS +6	+4 +5	+3 −2	−3 −6	
	%	%	%	%	
OSS	33	18	12	6	
S	55	56	55	31	
F	12	26	33	63	
Total No.	42	45	51	35	173

Age 30 and Over

RATINGS BY MANAGERS	SCORES IN INTEREST ANALYSIS +6	+4 +5	+3 −2	−3 −6	
	%	%	%	%	
OSS	29	24	13	7	
S	53	50	52	43	
F	18	26	35	50	
Total No.	58	42	71	44	215

In the ages 24 and under we get one variation from increase in failures with decrease in score. This may be simply a chance variation since the number of persons in the age score group is only 42 or it may actually have some significance. If it has significance it may be that overly much centering of interest in the younger ages is not quite so advantageous. We shall check this with future results.

One item of interest is that of the 19 who were under 25 and who scored a minus 6, there were 17 failures and 2 probable failures.

The group having the best possible chance of success are those from 25–29 who have a high centering of interest around the sales group represented by a plus 6. Here, as will be noted, out of 42 persons, 88% were successful. On the other hand, if one is going to contract with a low score man, it is best to do so with one 30 or over, as here only 50% are unsuccessful

as against 75% of those 24 and under and 63% of those from 25–29.

We have considered those who returned to family agencies in a separate group, since, as might be expected, the relative chance of success is higher. Since there were only 99 in the total group we condensed the tables to ratings of "success" and "failure" only, to get a sufficient number for percentages to have meaning. In the case of only 27 who were over 24 but in the fathers' agencies we have condensed the number of score groups to three. It is interesting, however, to note that even under this unfavorable statistical condition there is still a definite and consistent tendency for a lower score to increase the percentage of failure and decrease the percentage of success.

24 and Under—Father's Agency

RATINGS BY MANAGERS	SCORES IN INTEREST ANALYSIS				
	+6	+4 +5	+3 −2	−3 −6	
	%	%	%	%	
S	83	75	68	59	
F	17	25	32	41	
Total No.	18	16	16	22	72

25 and Over—Father's Agency

RATINGS BY MANAGERS	SCORES IN INTEREST ANALYSIS			
	+6	+2 −2	−3 −6	
	%	%	%	
S	100	57	40	
F	0	43	60	
Total No.	8	14	5	27

Two items made us feel that the data have definite meaning. The first is the internal consistency of the data. Almost without exception, the lowering of scores lowers the percentage of both outstanding successes and successes and increases the per-

centage of failures. The second item is more vague than this, but comes from having read the memoranda concerning the various individuals that fall out of line with the test. For example, one of the three outstanding successes of men over 30 who had a score of minus 6 was a man whose chief interest has always been engineering and whose father-in-law bought him an agency in order to have his daughter near home. The man is making an outstanding success of it, but at the last report, without enthusiasm for the work. Several of those who scored high but were failures are reported as having too much money and not needing to work.

The Ætna Casualty and Surety School group was an ideal group upon which to experiment since the situation is very much that which one gets in all selection; that is, a pre-selection by the managers or Home Office representatives previous to the giving of the test material. While this makes differentiation more difficult for the test it does duplicate the conditions under which tests can be most useful; that is, where they are used as an aid to a careful selection by interview and reference.

A FURTHER ANALYSIS OF THE YOUNG–ESTA–BROOKS STUDIOUSNESS SCALE

E. G. WILLIAMSON
University of Minnesota

An earlier study involving Arts College freshmen in the University of Minnesota indicated limited validity for the Young-Estabrooks Studiousness Index.[1] Because of the small sampling of students involved in this first study, a further analysis was made of the Studiousness scores of 539 freshmen men enrolling in the Arts College in September of 1936.

The correlation between Studiousness and first quarter honor point ratio was .20; and .10 with the American Council on Education Psychological Examination. The A. C. E. test correlated with grades to the extent of .45. The combination of A. C. E. and Studiousness scores yielded a multiple R with grades of .48, a negligible increase over A. C. E. alone. Moreover, the correlation of high school grades with college grades (.56) is not increased at all by the addition of Studiousness scores.

It is apparent, therefore, that the Studiousness Scale has negligible validity for Arts College freshmen in the University of Minnesota, a conclusion supported by analysis of the records of three different groups of students. That this condition is not true for other colleges is indicated by reports from Colgate University and the University of Florida.[2] Since the distribution of scores on the Studiousness Scale is similar at Colgate and Minnesota, differences in validity coefficients are not caused by this factor. It is possible that the criterion of scholarship at Minnesota differs to such an extent from that of Colgate and Florida that the Studiousness Scale is not applicable at Minnesota. It is also likely that the low odd-even reliability coefficient of .8054 (doubled) derived from the test scores of the 539 men in this study may have affected the above validity coefficients. Whatever the causes of these coefficients, it is apparent that similar research should be conducted at other colleges before the Scale is used to predict scholarship or to counsel students.

[1] Williamson, E. G., "An Analysis of the Young-Estabrooks Studiousness Scale," *The Journal of Applied Psychology*, June, 1937, pp. 260–264.

[2] Young, C. W. and Estabrooks, G. W., "Report on the Young-Estabrooks Studiousness Scale for Use with the Strong's Vocational Interest Blank for Men," *Journal of Educational Psychology*, Vol. XXVIII, No. 3, March, 1937, pp. 176–188.

Mosier, Charles I., "Factors Influencing the Validity of a Scholastic Interest Test," *Journal of Educational Psychology*, Vol. XXVIII, No. 3, March, 1937, pp. 188–196.

NEWS AND NOTES

The Midwestern Psychological Association will hold its annual meeting at the University of Wisconsin, Madison, Wisconsin, on April 22 and 23, 1938, under the presidency of Dr. Arthur G. Bills of the University of Cincinnati. The title of Dr. Bills' presidential address will be "Changing Views of Psychology as Science." A special feature of the meeting will be the celebration of the fiftieth anniversary of the founding of the psychological laboratory at the University of Wisconsin, one of America's earliest, by Dr. Joseph Jastrow. It is hoped that both Dr. Jastrow and Dr. Clark L. Hull, now of Yale University but formerly of the Department of Psychology at the University of Wisconsin, will be present at this fiftieth anniversary of the founding of the psychological laboratory.

Nearly 800 children are using the services of the Indiana University psychological clinics at the school's medical center in Indianapolis and on the campus at Bloomington. Children are referred to the clinics by schools, social service agencies, physicians, juvenile courts and similar agencies. Remedial speech training, orthoptic training in cross-eye cases, psychological examinations and other measures are taken to help the young patients.

From Stanford University come the following announcements: Dr. Florence L. Goodenough, of the University of Minnesota, will join the teaching staff of that institution during the summer of 1938. Dr. Paul R. Farnsworth, of the faculty of Stanford University, will go to the University of Wisconsin for the coming summer session. Dr. Frank C. Davis, University of California at Los Angeles, will teach in the summer school of the University of British Columbia.

The adoption of a comprehensive charter for college personnel work and the reconstruction of its constitution were signal accomplishments of the American College Personnel Association at its annual meeting in Atlantic City, February 23 to 26. The constitution states that the purpose of the organization is to provide for the cooperative association of those persons engaged in student personnel service and its administration in colleges or universities, to promote national and regional conferences for

the discussion of personnel problems and issues, to formulate and to encourage personnel research, and to disseminate "personnel point of view" among all educational workers. Membership in the association is open to all persons "engaged in any aspect of non-profit-making student personnel service, administration, or research at the college level" and also to a limited number of personnel workers in business and industry.

The program for American Education Week 1938 has just been announced by the National Education Association. This program has been adopted by the three national agencies which sponsor American Education Week—the National Education Association, the United States Office of Education, and the American Legion.

General Theme: *Education for Tomorrow's America.*

Sunday, November 6	Achieving the Golden Rule
Monday, November 7	Developing Strong Bodies and Able Minds
Tuesday, November 8	Mastering Skills and Knowledge
Wednesday, November 9	Attaining Values and Standards
Thursday, November 10	Accepting New Civic Responsibilities
Friday, November 11	Holding Fast to Our Ideals of Freedom
Saturday, November 12	Gaining Security for All.

EDUCATIONAL ABSTRACTS, formerly published by Norman J. Powell and Associates of New York City, was formally accepted by the National Council of Phi Delta Kappa, professional and fraternal association of men in education, at its biennial meeting in Cincinnati, December 29, 1937. The November-December number of the magazine was published by the new owner. Paul M. Cook has been named to the editorship of the magazine, with W. A. Stumpf as associate editor; a staff of seventy-five abstractors prepares the material for publication. Editorial offices are located at 1180 East Sixty-third Street, Chicago, Illinois.

Secretary of the Interior, Harold L. Ickes, recently announced the appointment of Dr. Helen K. Mackintosh, of Oxford, Ohio, to the position of Senior Specialist in Elementary Education in the Office of Education. Miss Mackintosh brings to her new position an exceptional background, particularly in the elementary school field, having earned A.B., M.A. and Ph.D., degrees at the State University of Iowa. She has had a broad and varied educational experience as a high school teacher, elementary school teacher and supervisor, and university professor.

"Social Education" will be the theme of the 1938 Stanford Education Conference, to be held at Stanford University, California, July 6–10. Forum sessions during the conference will be devoted to discussion of

experiments, investigations, and programs in social education and social control; to appraisal of practices and trends in the field; and to interpretation of the educational implications of American culture. Among the speakers to take part in the program are William Heard Kilpatrick, Emeritus Professor of Education, Columbia University; Lewis Mumford, author and lecturer; William Ogburn, University of Chicago; and Ray Lyman Wilbur, President of Stanford University.

The Annual Report of the Secretary and Director of Research of the Committee for the Study of Suicide, Inc., has recently been submitted to the members and board of directors. The activities of this Committee during the past year have been concentrated on clinical work. (1) Dr. John Benjamin, under the direction of Dr. Franklin G. Ebaugh of the Psychopathic Hospital of the University of Colorado, continued his psychoanalytic studies of fifty suicidal patients and also his observations in connection with the Rorschach test. Thus far this test seems to suggest that differential diagnostic criteria might be established between suicidal fantasies which remain fantasies and those developing into suicidal drives. (2) An intensive psychoanalytic study of six suicidal patients, carried on by members of the staff of the Institute for Psychoanalysis in Chicago, will no doubt show the value of this method of therapy in suicidal problems. (3) The work at Bellevue Hospital, New York City, is much more extensive as to numbers. Out of 1354 patients seen, 254 have been studied, of which approximately 27 were studied in detail from the standpoint of psychological and social background as well as of the developmental history of the mental illness. (4) At McLean Hospital, Waverley, Mass., of 154 patients available for study, thirty have made suicidal attempts and twenty-six suicidal threats. Dr. M. Ralph Kaufman has begun psychoanalysis of these patients and the collection of all possible data on family background. (5) Research at the Payne Whitney Psychiatric Clinic of the New York Hospital has been carried on since July 1, 1937, by Dr. Valer Barbu. He is attempting to evaluate the determinants of the suicidal drives in patients who have strong suicidal tendencies or who have made suicidal attempts; also what factors are at work which could be considered favorable or unfavorable as far as response to treatment is concerned, and what criteria for safety are to be considered when such patients are about to be discharged from the hospital. (6) A study of suicide among the Mohave Indians by Dr. George Devereux, formerly of the University of California, has been completed. He is at present organizing his material on suicide among the Moi in Indo-China.

The Board of Directors and officers of the Committee for the Study of Suicide include Dr. Gerald R. Jameison, President; Mr. Marshall Field,

Vice-President; Dr. Henry Alsop Riley, Treasurer; Dr. Gregory Zilboorg, Secretary; Miss Elizabeth G. Brockett, Dr. Franklin G. Ebaugh, Dr. Herman Nunbert, Dr. Dudley D. Schoenfeld and Dr. Bettina Warburg. The Executive Offices are located at Room 1404, Medical Arts Center, 57 West 57th Street, New York City.

The H. W. Wilson Company, New York City, recently announced the publication of a book entitled "My Vocation: By Eminent Americans," and arranged by Earl B. Lockhart. Contributors include such leading Americans as Dr. John Dewey, Dr. Harry Emerson Fosdick, Rear Admiral Richard E. Byrd, Professor George P. Baker, of Yale University, Judge Florence A. Allen, United States Court of Appeals, and the late Jane Addams of Hull House fame. Excellently reproduced full page portraits and biographical sketches of these distinguished contributors personalize each of the essays.

The Sesquicentennial edition of the famous FEDERALIST papers has been published by the National Home Library Foundation of Washington, D. C., as a part of the celebration of the one hundred fiftieth anniversary of the Constitution. THE FEDERALIST essays were originally written in 1787 by Alexander Hamilton, James Madison and John Jay as an attempt to explain the Constitution to the people of the United States. This edition has been prepared by several of the most distinguished American historians under the editorship of Professor Edward Mead Earle of Princeton University. It is issued on a non-profit basis and is nominally priced—paper edition 50 cents, cloth edition, 75 cents, and gift edition $1.00. It contains the complete collection of the eighty-five essays, the Articles of Confederation, The Constitution of the United States, and an index, a total of 675 pages.

The American Association on Mental Deficiency is to hold its sixty-second annual meeting in Richmond, Virginia, April 20–23, 1938, with headquarters at Hotel Jefferson. The officers of the Association for 1937–38 are: President, Harry C. Storrs, M.D., Superintendent, Letchworth Village, Thiells, New York; Vice-President, Neil A. Dayton, M.D., State Department of Mental Diseases, Boston, Mass.; Secretary-Treasurer, E. Arthur Whitney, M.D., Superintendent, Elwyn Training School, Elwyn, Pa. The Committees on Program and Arrangements have provided full and interesting programs for the various sessions. Among the many outstanding names appearing on the programs are Dr. Fred Kuhlmann, of Minnesota; Dr. Edgar A. Doll, Vineland, N. J.; Elaine F. Kinder, Letchworth Village, Dr. Thorlief G. Hegge, Northville, Michigan; Lloyd N. Yepsen, Trenton, N. J. At the Thursday morning Breakfast for the Ladies, Mrs. Eleanor Roosevelt will be the guest of honor.

BOOK REVIEWS

The Education of Personality. M. M. SUFIAN, Published by B. Ahmed, Derbar Press, Calcutta, 1937.

This little book maps out an ambitious project, one which, in my opinion, is exceedingly important: the nature-nurture controversy in relation to social theory. The author ranges the whole field of psychology from the extremely biological problems of Weissmanism to Marxian economics.

Unfortunately, he lacks the necessary training for his task. His chapter on inheritance, for example, is sadly muddled. He discusses Mendel, Weissman and Terman as exponents of "the inheritance of acquired characters." More or less intuitively, however, he reaches a conclusion which most American psychologists today would accept—*i.e.*, that environment is at least as important as heredity in determining intellectual achievement.

His thinking on the relation of heredity to social theory also runs into difficulties. *E.g.*, at one time he protests an emphasis on heredity as anti-democratic (p. 12) and very correctly points to Hitler as an apostle of an extreme type of thinking on heredity. But in another connection (p. 54) he says that the industrial revolution made democracy necessary so that worthy individuals could rise to places of power in the economic world—thus bowing to the inevitability of innate differences.

The truth of the matter seems to be, not that our view of society is shaped by our position on the nature-nurture controversy, but rather that our social theory will determine the application we make of ambiguous facts regarding heredity and environment. This is a problem which could do with much more study, particularly by American psychologists, who are prone to bury their heads in statistical sands to avoid looking at social problems.

ROSS STAGNER,
University of Akron.

FRED S. KELLER. *The Definition of Psychology.* New York, Appleton-Century Company, 1937. Pp. vi + 111. $1.00.

This book accomplishes what it was intended that it should accomplish. It leads to an adequate definition of what psychology is in terms of what psychologists do, and the material is presented in such a way that college sophomores can understand it.

The approach to the definition is historical and touches upon the most

important events and people. The first chapter deals with the contributions of Descartes, the British empiricists and Fechner.

The second chapter is concerned with the "Founding of Modern Psychology" and presents the contributions of Wundt, dealing with the subject matter, the methods and the problems of psychology. From this point the author rounds out the treatment of Structuralism with the expansion, extension and refinement that Titchener gave it. He follows this by a chapter on each of Functionalism, Behaviorism and Gestalt psychology relating these schools to one another and to Structuralism and pointing out the contributions of the leaders in each school. Finally, it is suggested that psychology is becoming more unified rather than more diversified. Further, that psychology is more objective, stressing behavior rather than experience. It is probable that most readers will agree with the author on the following points: that psychology is becoming more separated from physiology and depending more upon its own set of facts; that it is turning more to observation of longi-sections of behavior rather than cross-sections; that it is more genetic and comparative than formerly.

The text of this book covers 106 pages and is followed by references which may be consulted by those students who wish to go more completely into the subject. Undoubtedly, this is simplifying the subject and a great many details are omitted, but what is presented here makes a complete story in itself and makes it available to many more students than the previous, more detailed histories of psychology.

K. W. Oberlin,
University of Delaware.

Harold E. Burtt. *Psychology of Advertising.* New York: Houghton Mifflin Co., 1938. Pp. x + 473.

A critical summary of the more recent experimental literature on the psychology of advertising has been lacking, but now this necessity has been met. Dr. Burtt, due to his experience in teaching and experimentation in the field was the logical person to write such a text.

As it is so well stated in the preface of the book, the material contained in the text should be intelligible not only to the student, but also to "the advertising executive, the copy-writer, or the business man who prepares his own advertisements." The author might well have added that *anyone* interested in having a thorough background in advertising psychology would find the text indispensable.

In general, the chapter titles follow those of former books on the psychology of advertising, including topics such as the prospect's wants, instincts and desires, long-circuit appeals, attention and size, position of the advertisements, color, and the like. A notable addition, however, includes a chapter on radio as an advertising medium.

Dr. Burtt's style makes for easy reading, and the careful outlining of the chapters and sub-topics allows one to follow the author's trend of thought easily. The topics are very extensive, and are dealt with fairly intensively, with one exception. It appears that Dr. Burtt might well have spent more space discussing the methods of testing advertisements, and a more critical attitude might well have been maintained here, inasmuch as all the generalities about the psychology of advertising are ultimately dependent upon the validity of the testing methods used.

Nevertheless, the book is an unquestionable contribution to the field, and the highest tribute one can pay to Dr. Burtt is to state that the work admirably rounds out his texts on the psychology of business, being a worthy complement to his *Principles of Employment Psychology* and *Psychology and Industrial Efficiency*.

ALBERT B. BLANKENSHIP,
Psychological Corporation.

SEASHORE, C. E. (Ed.). *Objective Analysis of Musical Performance: Studies in the Psychology of Music.* Volume IV. University of Iowa Studies, Iowa City, 1936, pp. 379.

This fourth volume in the University of Iowa Studies in the Psychology of Music contains ten papers devoted to analyzing and measuring the musical performance of artists in an attempt to understand their artistry and the nature of musical esthetics. Eventually it is hoped that further studies in this field will disclose the main laws of musical esthetics and that norms of artistic performance will be established.

The paper by Harold C. Seashore on "An Objective Analysis of Artistic Singing" is the model for investigations in this whole field, and is an ambitious attempt to measure the most important factors in artistic singing and then to establish norms for each of these factors of artistic performance. These norms are based on different aspects of sound—pitch, intensity, duration. Rhythm and phrasing as additional factors are too complex for normative treatment but are analyzed into their component elements and are related to performance. The study is limited to an analysis of ten songs of the legato type so the results are not necessarily applicable to all musical forms such as the sherzo, the waltz, or all operatic forms. Five of the songs are Victor recordings and five are direct phonophotographic records of the artists themselves. Perhaps the most important conclusion is that artistry is not to be measured by any fixed mechanical standard but is a matter of subtle variations from a standard. This principle of "orderly deviations from the regular" (p. 98) holds for all factors considered, thus: (1) performance as compared to strict notation, (2) one performer as compared to another, (3) the same performer compared to his own later rendition of the same piece,

(4) absolute standards as compared to performance standards in pitch or duration, and (5) other factors such as vibrato, intensity, rhythm, and phrasing. Although this principle is by no means new, it is valuable to have objective examples of its application and generality.

Ray S. Miller studies "The Pitch of the Attack in Singing" in ten Victor artist records and in six amateurs. Fifty-eight per cent of all attacks were rising glides, or in other words, the beginning of the tone is low and gradually rises as the tone is held. The amount of the glide depends upon several factors.

"An Objective Analysis of Artistic Violin Performance" by Arnold M. Small does the same thing for violin playing that Harold Seashore's investigation did for singing and with almost identical results, showing that the factors involved are, with the exception of intonation and extent of vibrato on transitions from note to note, fairly constant for the two types of performances.

Paul C. Greene analyzes "Violin Performance with Reference to Tempered, Natural and Pythagorean Intonation" and finds that six violinists tend to follow the Pythagorean scale in intonation rather than the natural or equally tempered scales. They also tend to sharp major seconds and thirds, and to flat minor seconds and thirds.

"The Iowa Piano Camera and Its Use" is described by M. T. Henderson, J. Tiffin, and C. E. Seashore. It is so constructed that it records the intensity and duration of every note on the piano. A transcription of this record into a "musical pattern score" makes possible the analysis of any piano performance at leisure, in detail, and quantitatively.

The first practical application of the Iowa Piano Camera is reported in a study by Laila Skinner and C. E. Seashore, "A Musical Pattern Score of the First Movement of the Beethoven Sonata, Opus 27, No. 2." The importance of this and similar studies to the musician can be seen from the following quotation, "He can identify marks of artistry, verify art principles, discover data for new principles, and characterize the performance of the artist in great variety of detail in accordance with his interests. Artists may be compared; the results of specific piano techniques may be seen; criticism, both theoretical and practical, may have an objective basis. His problem can be defined; it can be isolated for study; the phenomenon in question may be reproduced; and the logic of science may operate in formulating the relevant facts for esthetic theory and practice." (p. 279).

M. T. Henderson contributes a second application of the Iowa Piano Camera in "Rhythmic Organization in Artistic Piano Performance" in which he shows that duration, rather than intensity, is the chief means of accentuating the first note in the measure of Chopin's Nocturne No. 6. Other techniques for studying rhythm are illustrated but the data are too meager to propose any principles.

A very complete study of "Synchronization in Artistic Piano Music" is capably handled by Leroy Ninde Vernon in which the data from four concert pianists—Bauer, Bachaus, Hofmann, Paderewski—indicate that chords may be played asynchronously to facilitate perception by bringing out the melody note in the chord, or to stress certain beats in the measure.

Finally, Don Lewis summarizes considerable data on pitch in an attempt to "make them meaningful for a psychological interpretation of music and speech." His article, "Pitch: Its Definition and Physical Determinants" brings out the fact that pitch is conditioned by intensity, duration, and harmony, as well as by frequency of the sound wave and suggests the need for further research with these factors.

The apparatus and techniques contributed by the Iowa Psychological Laboratory open up the possibility of an exact quantitative treatment of the principles of musical esthetics which may be of great practical importance to the musician. The articles in the present volume are pioneers in this field which illustrate what can be done in the field of art. The musician of the future will have to be less reckless with his theorizing and dogmatism for every principle of music appreciation or performance can be objectively tested.

RALEIGH M. DRAKE,
Wesleyan College,
Macon, Georgia.

Logic, Theoretical and Applied. By D. LUTHER EVANS AND WALTER S. GAMERTSFELDER. Garden City. Doubleday, Doran & Company, 1937. Pp. xiii, 482.

The present text is primarily a restatement of the traditional logic as seen through the eyes of the authors, preceded by a section on The Nature and Value of Reflective Thought, the intention of which is to give formal principles a setting in the general field of knowledge, to suggest the relation of logic to language, the sciences, and philosophy; and also to relate it to the problem of truth.

The discussion of the Nature and Function of Terms has in it nothing beyond the traditional presentation. In the treatment of the proposition recognition is given to relational forms, propositional functions, primitive propositions, etc. But these do not form an integral part of the treatment. They stand as added thoughts, with no fundamental bearing upon the heart of the discussion.

The treatment of immediate inferences as cases of *equivalence* is incorrect. Certainly conversion and inversion do not come under this description; nor do those instances in which no immediate inference is possible! The problem of existential import receives no consideration in spite of the fact that it has been the subject of much discussion and must be given a

certain interpretation in order to make the square of opposition acceptable.

I often wonder why writers of elementary logics do not recognize that the alternative syllogism is not a simple form of argument, but involves, when analyzed, the principles of "composition," "distribution," and "simplification," now commonly recognized by symbolic logic, and also the categorical syllogism. The ancient tradition was right in treating the categorical syllogism first.

The discussion of the categorical syllogism follows the customary procedure, speaking as if the premises and conclusion are always given before the argument is discussed. This throws the syllogism open to the charge that it is purely a logic for proof, not a basis for discovery or anticipation. Why not regard the premises as data formulated, as the basis for a forward movement of thought to the conclusion. From this point of view the categorical syllogism becomes as much a matter of anticipation as does an inductive procedure. The definitions of terms should be altered to correspond, the minor becoming the topic about which something is to be asserted, the major the term in which it is to be expressed. Traditional logic is still valid. It is constantly used. Our ways of approaching and discussing it are what bring it into disfavor, by making it apparently artificial and trivial. The attempt to save it by a metaphysical discussion of The Philosophical Import of the Syllogism is futile. Elementary students of logic have no philosophic background.

The chapter on Irregularities of Syllogistic Thought includes relational arguments along with the traditional prosyllogism, etc. It is time these were not judged from the syllogistic point of view and regarded as "irregular."

The chapter on the Meaning of Science is better. It does more justice to the various meanings which have been given to "induction." And Mill's methods on causation are rightly placed in a separate chapter. Chapters on Analogy, Probability, and Fallacies are in the traditional vein. The discussion of Logic and the specialized sciences will not serve very effectively to connect the two in the minds of students, I suspect. The chapter on values is more successful, though it has much which goes a good distance beyond logic.

As a whole the book is as good as any of the traditional presentations, and in its mode of stating things will doubtless appeal more to certain teachers than other texts do. But students do not gain a fair notion of what logic is doing at the present day from a mere restatement of the traditional discussions.

ALBERT E. AVEY,
Ohio State University.

NEW BOOKS AND PAMPHLETS RECEIVED

Books and pamphlets for review should be sent to JAMES P. PORTER, *Editor*, JOURNAL OF APPLIED PSYCHOLOGY, Ohio University, Athens, Ohio.

Abnormal Psychology. Students Outline Series. RALEIGH M. DRAKE. Longmans, Green & Co., New York, 1937. 75 cents. 136 pp.

A Biological Approach to the Problem of Abnormal Behavior. MILTON HARRINGTON. Science Press Printing Company, Lancaster, Pa., 1938.

The Case for Democracy. ORDWAY TEAD. Association Press, New York, 1938. $1.25. 120 pp.

De Busk Memorial Essays. HENRY DAVIDSON SHELDON, *Editor.* University of Oregon, Eugene, Oregon. 68 pp.

Educational Tests and Measurements. Students Outline Series. WENDELL W. WRIGHT. Longmans, Green & Company, New York, 1937. 75 cents. 99 pp.

Finding Your Work. J. GUSTAV WHITE. Association Press, New York, 1938. 64 pp.

A First Course in Statistics. E. F. LINDQUIST. Houghton Mifflin Company, Boston, Mass., 1938. $2.25. 226 pp.

General Psychology. WILLIAM STERN. Macmillan Company, New York, 1938. $4.50. 589 pp.

A Handbook for Testing Extra-Sensory Perception (With ESP Cards). C. E. STEWART AND J. G. PRATT. Farrar & Rinehart, New York, 1937. $1.75. 96 pp.

How to Teach. CLAUDE C. CRAWFORD. Southern California Book Depository, Los Angeles, Calif., 1938. $2.50. 511 pp.

The Intelligence Quotient: Is It a Constant? CLARA HARRISON TOWN. Children's Aid Society, 70 W. Chippewa Street, Buffalo, N. Y. 25 cents. 38 pp.

An Introduction to the Fields of Psychology. EMILY S. DEXTER AND KATHARINE T. OMWAKE. Prentice-Hall, New York, 1938. $1.75. 236 pp.

Manual of Psychiatry and Mental Hygiene. Seventh Edition. AARON J. ROSANOFF. John Wiley & Sons, New York, 1938. $7.50. 1091 pp.

Mind in Transition. JOSEPH K. HART. Covici-Friede, New York, 1938. $3.50. 413 pp.

The Modern Mind. MICHAEL ROBERTS. Macmillan Company, New York, 1937. $2.75. 284 pp.

New Frontiers of the Mind. J. B. RHINE. Farrar & Rinehart, New York, 1937. $2.50. 275 pp.

Personnel and Labor Relations. DALE YODER. Prentice-Hall, New York, 1938. $5.35. 664 pp.

The Problem of Education. CLAUDE C. CRAWFORD, LOUIS P. THORPE AND FAY ADAMS. Southern California School Book Depository, Los Angeles, Calif., 1938. $2.50. 239 pp.

Psychology of Advertising. HAROLD ERNEST BURTT. Houghton Mifflin Company, Boston, Mass., 1938. $3.00. 473 pp.

Psychology. Students Outline Series. KATE HEVNER. Longmans, Green & Company, New York, 1937. 75 cents. 96 pp.

Radio: A Discussion for Executives. H. K. BOICE, Columbia Broadcasting System. (A chapter from *Handbook of Advertising,* Weiss, Larabee and Kendall) McGraw-Hill Book Company, New York. 49 pp.

Reminiscence and Rote Learning. LEWIS B. WARD. Psychological Monograph, Vol. 49, No. 4. Psychological Review Company, Ohio State University, Columbus, Ohio, 1938. 64 pp.

Social Psychology of Education. A. O. BOWDEN AND IRVING R. MELBO. McGraw-Hill Book Company, New York, 1937. $2.50. 296 pp.

Statistical Reference Data Relating to Problems of Overageness, Educational Retardation, Non-Promotion, 1900–34, Part II. EUGEN A. NIFENECKER. Board of Education, New York City, Bureau of Reference, Research and Statistics, July 1937. 287 pp.

Student Guidance Techniques. D. G. PATERSON, G. G. SCHNEIDLER AND E. G. WILLIAMSON. McGraw-Hill Book Company, New York, 1938. $3.00. 316 pp.

The Study of Human Abilities. American Oriental Studies, Vol. II. W. NORMAN BROWN, JOHN K. SHRYOCK AND E. A. SPEISER. American Oriental Society, New Haven, Conn., 1937. 168 pp.

Study Manual for a First Course in Statistics. E. F. LINDQUIST. Houghton Mifflin Company, Boston, Mass., 1938. 80 cents. 122 pp.

Directory of American Psychological Periodicals

(*Continued*)

Comparative Psychology Monographs—Baltimore: Johns Hopkins University Press. Subscription $5.00. 400 pages per volume. Knight Dunlap, Managing Editor. Published without fixed dates, each number a single research. Founded 1922.

Genetic Psychology Monographs—Provincetown, Mass.: The Journal Press. Subscription $7.00 per volume. One volume per year. Edited by Carl Murchison and an international cooperating board. Quarterly. Each number one complete research. Child behavior, animal behavior and comparative psychology. Founded 1925.

Psychological Abstracts—Eno Hall, Princeton, N. J.: American Psychological Association. Subscription $6.00. 700 pages ann. Edited by W. S. Hunter and R. R. Willoughby. Monthly. Abstracts of psychological literature. Founded 1927.

Journal of Genetic Psychology—Provincetown, Mass.: The Journal Press. Subscription $14.00 per annum; $7.00 per volume. 1,000 pages annually (2 volumes). Edited by Carl Murchison and an international cooperating board. Quarterly. Child behavior, animal behavior, and comparative psychology. Founded 1891.

Journal of Social Psychology—Provincetown, Mass.: The Journal Press. Subscription $7.00. 500–600 pages annually. Edited by John Dewey, Carl Murchison, and an international cooperating board. Quarterly. Political, racial, and differential psychology. Founded 1929.

Personnel Journal—Baltimore: The Williams & Wilkins Co. Subscription $5.00. 500 pages, approximately. Edited by Walter V. Bingham. Bi-monthly.

Mental Measurement Monographs—Baltimore: The Williams & Wilkins Company. Subscription $5.00. 400 pages annually. Edited by Knight Dunlap and Buford Johnson. Published without fixed dates, each number a single research. Founded 1925.

Studies in Psychology and Psychiatry—Baltimore: The Williams & Wilkins Company. Subscription $5.00. 400 pages annually. Edited by Edward A. Pace. Published without fixed dates, each number a single research. Founded 1926.

Journal of Educational Psychology—Baltimore, Md.: Warwick & York. Subscription $6.00. Edited by Jack W. Dunlap, Harold E. Jones, and Percival M. Symonds. Published monthly except June to August. Founded 1910.

Child Development—Baltimore, Md.: Williams & Wilkins Company. Subscription $5.00. 400 pages annually. Edited by Buford Johnson. Quarterly. Founded 1930.

Character and Personality—Durham, N. C.: Duke University Press. Subscription $2.00 per annum. Quarterly. Published simultaneously in Great Britain and the United States. Edited by Charles E. Spearman (London) and Karl Zener (Duke University). Founded 1932.

Journal of Psychology—Provincetown, Mass.: The Journal Press. Subscription $14.00 per annum; $7.00 per volume. 1000–1200 pages annually. Immediate publication for any accepted manuscript. Edited by Carl Murchison and an editorial board. Quarterly. The general field of psychology. Founded 1936.

Directory of American Psychological Periodicals

American Journal of Psychology—Ithaca, N. Y.: Cornell University. Subscription $6.50. 624 pages annually. Edited by K. M. Dallenbach, M. F. Washburn, M. Bentley, E. G. Boring. Quarterly. General and experimental psychology. Founded 1887.

Journal of General Psychology—Provincetown, Mass.: The Journal Press. Subscription $14.00 per annum; $7.00 per volume. 1000 pages annually (2 volumes). Edited by Carl Murchison and an international cooperating board. Quarterly. Experimental, theoretical, clinical, and historical psychology. Founded 1927.

Psychological Review—Princeton, N. J.: Psychological Review Company. Subscription $5.00. 540 pages annually. Bi-monthly. General Psychology. Edited by Herbert S. Langfeld. Founded 1894.

Psychological Monograph—Princeton, N. J.: Psychological Review Company. Subscription $6.00 per vol. 500 pp. Ed. by Joseph Peterson. Published without fixed dates, each issue one or more researches. Founded 1895.

Psychological Index—Princeton, N. J.: Psychological Review Company. Subscription $4.00. 300–400 pp. Edited by W. S. Hunter and R. S. Willoughby. An annual bibliography of psychological literature. Founded 1895.

Psychological Bulletin—Princeton, N. J.: Psychological Review Company. Subscription $6.00. 720 pages annually. Psychological literature. Monthly (10 numbers). Edited by John A. McGeoch. Founded 1904.

Archives of Psychology—New York: Columbia University. Subscription $6.00. 500 pp. ann. Ed. by R. S. Woodworth. Published without fixed dates, each number a single experimental study. Founded 1906.

Journal of Abnormal and Social Psychology—Eno Hall, Princeton, N. J.: American Psychological Association. Subscription $5.00. Foreign, $5.25. Editor, Harry T. Moore. Quarterly. 448 pages annually. Abnormal and Social. Founded 1906.

Psychological Clinic—Philadelphia: Psychological Clinic Press. Subscription $3.00. 288 pages. Ed. by Lightner Witmer. Without fixed dates (Quarterly). Orthogenics, psychology, hygiene. Founded 1907.

Psychoanalytic Review—Washington, D. C.: 3617 10th St., N.W. Subscription $6.00. 500 pages annually. Psychoanalysis. Quarterly. Ed. by W. A. White and S. E. Jelliffe. Founded 1913.

Journal of Experimental Psychology—Princeton, N. J. Psychological Review Company. 500 pages annually. Experimental. Subscription $7.00. Bi-monthly. Edited by S. W. Fernberger. Founded 1916.

Journal of Applied Psychology—Lancaster, Pa.: Science Press Printing Co. Subscription $6.00. 500 pages annually. Bi-monthly. Edited by James P. Porter, Ohio University, Athens, Ohio. Founded 1917.

Journal of Comparative Psychology—Baltimore: The Williams & Wilkins Co. Subscription $5.00 per volume of 450 pages. Two volumes every year. Edited by Knight Dunlap and Robert M. Yerkes. Founded 1921.

(Continued on Third Cover)